(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 December 2003 (04.12.2003)

PCT

(10) International Publication Number
WO 03/099865 A1

(51) International Patent Classification[7]: C07K 14/705, C12N 5/10, 15/12, C12Q 1/68, A01K 67/027, A61K 38/02, 39/00, G01N 33/48, 33/50

(21) International Application Number: PCT/GB03/02270

(22) International Filing Date: 23 May 2003 (23.05.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0212067.3 24 May 2002 (24.05.2002) GB

(71) Applicant *(for all designated States except US)*: **INPHARMATICA LIMITED** [GB/GB]; 60 Charlotte Street, London W1T 2NU (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LOBLEY, Anna, Elizabeth** [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). **MICHALOVICH, David** [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). **ALLEN, Kathryn, Elizabeth** [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). **REYNOLDS, Lindsey** [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). **PIERRON, Valerie, Nathalie** [FR/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB).

(74) Agents: **GOODFELLOW, Hugh, Robin** et al.; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/099865 A1

(54) Title: CATION CHANNEL PROTEINS

(57) Abstract: This invention relates to novel proteins (termed INPIONCH05 and INPIONCH06), herein identified as cation channels and to the use of these proteins and nucleic acid sequences from the encoding genes in the diagnosis, prevention and treatment of disease.

CATION CHANNEL PROTEINS

This invention relates to novel proteins (termed INPIONCH05 and INPIONCH06), herein identified as cation channels and to the use of these proteins and nucleic acid sequences from the encoding genes in the diagnosis, prevention and treatment of disease.

All publications, patents and patent applications cited herein are incorporated in full by reference.

BACKGROUND

The process of drug discovery is presently undergoing a fundamental revolution as the era of functional genomics comes of age. The term "functional genomics" applies to an approach utilising bioinformatics tools to ascribe function to protein sequences of interest. Such tools are becoming increasingly necessary as the speed of generation of sequence data is rapidly outpacing the ability of research laboratories to assign functions to these protein sequences.

As bioinformatics tools increase in potency and in accuracy, these tools are rapidly replacing the conventional techniques of biochemical characterisation. Indeed, the advanced bioinformatics tools used in identifying the present invention are now capable of outputting results in which a high degree of confidence can be placed.

Various institutions and commercial organisations are examining sequence data as they become available and significant discoveries are being made on an on-going basis. However, there remains a continuing need to identify and characterise further genes and the polypeptides that they encode, as targets for research and for drug discovery.

CATSPER CATION CHANNELS

CatSper1 (Ren, D. *et al.* A sperm ion channel required for sperm motility and male fertility. Nature, 413(6856), 603-609 (2001); genbank accession:AAL14105) and CatSper2 (Quill,T.A. *et al.*, A voltage-gated ion channel expressed specifically in spermatozoa. PNAS, Vol. 98(22) 12527-12531 (2001); genbank accession: AAL26490) are related testis-specific transcripts with a single cationic transportation domain. They share 21% overall sequence identity and CatSper1 contains a longer N-terminal histidine rich portion that is absent in CatSper2.

A third type of CatSper channel has also recently been found within patent literature, which will be referred to here as CatSper3 (Lexicon genetics patent sequence: WO200066735-A2).

CatSper1 and CatSper2 are both testis-specific transcripts (Ren, D. *et al.* A sperm ion channel required for sperm motility and male fertility. Nature, 413(6856), 603-609 (2001); Quill,T.A. *et al.*, A voltage-gated ion channel expressed specifically in spermatozoa. PNAS, Vol. 98(22) 12527-12531 (2001)).

CatSper1 and CatSper2 are involved in sperm motility and are localised to the principal piece/flagellum of the tail region of sperm. The high abundance of histidine residues in the amino terminus of CatSper1 may be involved in the well-known pH regulation of sperm motility (Yanagimachi,R. The Physiology of Reproduction (eds. Knobill,E. & Neill,J.D.) 189-315 (Raven, New York, 1994); Sekler,I. *et al.* A cluster of cytoplasmic histidine residues specifies pH dependence of the AE2 plasma membrane anion exchanger. *Cell*, 86(6), 929-935 (1996)).

Targeted disruption of the CatSper1 gene has resulted in male sterility in otherwise normal mice (Ren, D. *et al).* A sperm ion channel required for sperm motility and male fertility. Nature, 413(6856), 603-609 (2001)) suggesting a potential role for members of the CatSper family in infertility. Ren *et al.* have found that although CatSper1 is not required for egg activation, it is required for the sperm to be able to penetrate the egg outer layers (zona pellucida).

Related ion channel families that feature the cationic transport domain that is present in CatSper1, CatSper2 and CatSper3 are the voltage-gated Na^+ and Ca^{2+} channel families. The alpha subunits of these channels all share a common 6 transmembrane-spanning domain that functions as the ionic selectivity filter and pore through the membrane (S5 and S6) and the voltage sensor (S4 domain) that is characterised by regular repeating positively charged residues within the transmembrane-spanning region. Both the voltage-gated Na^+ and voltage-gated T- and L-type Ca^{2+} channels contain a four domain repeat of the ion_trans domain.

Cationic sperm channels, transient receptor potential, polycystic kidney disease protein (PKD)/sea urchin receptor for egg jelly (REJ) receptors, cyclic nucleotide gated channels and voltage-gated potassium channels on the other hand all contain a single ion_trans domain. Cyclic nucleotide gated channels are believed to be tetramers composed of two homologous types of subunit, alpha and beta, whilst voltage-gated potassium channels are made up of four alpha subunits each containing a single ion_trans domain.

Although ion selectivity has not been determined, comparison of the amino acid sequences of CatSper1, CatSper2 and CatSper3 with other known channels shows that the pore forming

region most closely resembles a single, six-transmembrane-spanning repeat of the voltage-dependent Ca^{2+} channels. The likelihood that the CatSper channels are permeable to Ca^{2+} is further supported by the finding that the cyclic-AMP-induced Ca^{2+} influx is abolished in the sperm of CatSper1$^{-/-}$mice (Ren, D. *et al.* A sperm ion channel required for sperm motility and male fertility. Nature, 413(6856), 603-609 (2001)).

Patch clamp studies of human and mouse CatSper1 expressed in a variety of heterologous expression systems including *Xenopus* oocytes and HEK-293 cells have failed to detect a significant current resulting from changes in voltage, pH, osmolarity, and/or cyclic nucleotide concentration, when CatSper1 was expressed alone or in combination with expressed CNG β-channel subunits (CNG4 or CNG6) (Ren, D. *et al.* A sperm ion channel required for sperm motility and male fertility. Nature, 413(6856), 603-609 (2001)). Furthermore, no current was detected when CatSper2 was expressed either alone or in combination with CatSper1 or CNG channel β-subunits (Quill,T.A. *et al.* A voltage-gated ion channel expressed specifically in spermatozoa. PNAS, Vol. 98(22) 12527-12531 (2001)).

Immunoprecipitation experiments involving CatSper1 and CatSper2 expressed both alone and co-expressed together have failed to demonstrate that these two proteins form hetero-dimers (Quill,T.A. *et al.* A voltage-gated ion channel expressed specifically in spermatozoa. PNAS, Vol. 98(22) 12527-12531 (2001)).

There is strong evidence from other cationic channel family members to suggest that four copies of the ion_trans domain are required to form functional channel complexes. Since the majority of channels are comprised of tetramers, it is likely that the CatSper1 and/or CatSper2 subunits act as part of a tetrameric assembly together with other unknown or known channel sub-units, such as the above-mentioned CatSper3.

Alteration of CatSper channel activity is a means to alter the sperm phenotype and as such, identification of novel CatSper family channel subunits is highly relevant as these would also be likely to play a role in sperm motility and subsequent fertility and thus represent potential targets for the manufacture of fertility drugs or non-hormonal contraceptives for both men and women. For example, no specific target has yet been identified for the drug Nifedipine (Pfizer) which is used as a reversible contraceptive and acts by slowing the movement of Ca^{2+} through the cell membrane.

THE INVENTION

The invention is based on the discovery that the INPIONCH05 and INPIONCH06 proteins function as cation channels. INPIONCH05 encodes a 522 amino acid transcript on human chromosome 1p36.11, and contains a non-selective cation ion_trans domain plus flanking N- and C-terminal intracellular domains. INPIONCH06 encodes a mouse orthologue located on mouse chromosome 4.

INPIONCH05 and INPIONCH06 are cation channels. Sequence alignment of the pore region of each of INPIONCH05 and INPIONCH06 to the pore region of other voltage-gated channels shows that they are not K^+ channels. The multiple alignments of Figure 6 indicate that there is possible Ca^{2+}-specificity. Further support for Ca^{2+}-specificity comes from the observation that INPIONCH05 and INPIONCH06 both contain the TxDxW motif which is conserved in other CatSper channels which are known to be Ca^{2+}-selective (illustrated by knockouts, Ren, D. *et al.* A sperm ion channel required for sperm motility and male fertility. Nature, 413(6856), 603-609 (2001)). Thus, it is considered most likely that cation transport through INPIONCH05 and INPIONCH06 is specific for Ca^{2+}.

In the first aspect of the invention, there is provided a polypeptide which:

(i) comprises or consists of the amino acid sequence as recited in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO: 20 or SEQ ID NO:22;

(ii) is a fragment thereof having the activity of a polypeptide according to (i), or having an antigenic determinant in common with a polypeptide according to (i); or

(iii) is a functional equivalent of (i) or (ii).

By "the activity of a polypeptide according to (i)", we refer to activity characteristic of a voltage-gated cation channel and preferably activity characteristic of a member of the CatSper family of sperm-specific ion channels.

The polypeptide having the sequence recited in SEQ ID NO:2 is referred to hereafter as "the INPIONCH05 exon 1 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:4 is referred to hereafter as "the INPIONCH05 exon 2 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:6 is referred to hereafter as "the INPIONCH05 exon 3 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:8 is referred to hereafter as "the INPIONCH05 exon 4 polypeptide". The polypeptide having the sequence

recited in SEQ ID NO:10 is referred to hereafter as "the INPIONCH05 exon 5 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:12 is referred to hereafter as "the INPIONCH05 exon 6 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:14 is referred to hereafter as "the INPIONCH05 exon 7 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:16 is referred to hereafter as "the INPIONCH05 exon 8 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:18 is referred to hereafter as "the INPIONCH05 exon 9 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:20 is referred to hereafter as "the INPIONCH05 exon 10 polypeptide". Combining SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18 and SEQ ID NO:20 produces the sequence recited in SEQ ID NO:22. The polypeptide having the sequence recited in SEQ ID NO:22 is referred to hereafter as the INPIONCH05 polypeptide.

The term "INPIONCH05 exon polypeptides" as used herein includes polypeptides comprising or consisting of the INPIONCH05 exon 1 polypeptide, the INPIONCH05 exon 2 polypeptide, the INPIONCH05 exon 3 polypeptide, the INPIONCH05 exon 4 polypeptide, the INPIONCH05 exon 5 polypeptide, the INPIONCH05 exon 6 polypeptide, the INPIONCH05 exon 7 polypeptide, the INPIONCH05 exon 8 polypeptide, the INPIONCH05 exon 9 polypeptide, the INPIONCH05 exon 10 polypeptide or the INPIONCH05 polypeptide.

In one embodiment of the first aspect of the invention, the polypeptide according to this embodiment consists of the amino acid sequence recited in SEQ ID NO:22 or is a fragment or functional equivalent thereof. In a second embodiment, the polypeptide consists of the amino acid sequence recited in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO: 6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20 or SEQ ID NO:22, or is a variant thereof.

In a second aspect of the invention, there is provided a polypeptide which:

(i) comprises or consists of the amino acid sequence as recited in SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44 or SEQ ID NO:46;

(ii) is a fragment thereof having the activity of a polypeptide according to (i), or having an antigenic determinant in common with a polypeptide according to (i); or

(iii) is a functional equivalent of (i) or (ii).

By "the activity of a polypeptide according to (i)", we refer to activity characteristic of a voltage-gated cation channel and preferably activity characteristic of a member of the CatSper family of sperm-specific cation channels.

The polypeptide having the sequence recited in SEQ ID NO:24 is referred to hereafter as "the INPIONCH06 exon 1 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:26 is referred to hereafter as "the INPIONCH06 exon 2 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:28 is referred to hereafter as "the INPIONCH06 exon 3 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:30 is referred to hereafter as "the INPIONCH06 exon 4 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:32 is referred to hereafter as "the INPIONCH06 exon 5 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:34 is referred to hereafter as "the INPIONCH06 exon 6 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:36 is referred to hereafter as "the INPIONCH06 exon 7 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:38 is referred to hereafter as "the INPIONCH06 exon 8 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:40 is referred to hereafter as "the INPIONCH06 exon 9 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:42 is referred to hereafter as "the INPIONCH06 exon 10 polypeptide". The polypeptide having the sequence recited in SEQ ID NO:44 is referred to hereafter as "the INPIONCH06 exon 11 polypeptide". Combining SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42 and SEQ ID NO:44 produces the sequence recited in SEQ ID NO:46. The polypeptide having the sequence recited in SEQ ID NO:46 is referred to hereafter as the INPIONCH06 polypeptide.

The term "INPIONCH06 exon polypeptides" as used herein includes polypeptides comprising or consisting of the INPIONCH06 exon 1 polypeptide, the INPIONCH06 exon 2 polypeptide, the INPIONCH06 exon 3 polypeptide, the INPIONCH06 exon 4 polypeptide, the INPIONCH06 exon 5 polypeptide, the INPIONCH06 exon 6 polypeptide, the INPIONCH06 exon 7 polypeptide, the INPIONCH06 exon 8 polypeptide, the INPIONCH06 exon 9 polypeptide, the INPIONCH06 exon 10 polypeptide, the INPIONCH06 exon 11 polypeptide or the INPIONCH06 polypeptide.

The ENSEMBL partial predicted amino acid transcript ENST00000294438, annotated as a non-selective cation channel, also possesses sequence similarity to the CatSper channels mentioned above and is explicitly excluded from the scope of the first and second aspects of the present invention.

In one embodiment of the second aspect of the invention, the polypeptide consists of the amino acid sequence recited in SEQ ID NO:46 or is a fragment of or functional equivalent thereof. In a second embodiment, the polypeptide consists of the amino acid sequence recited in SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44 or SEQ ID NO:46, or is a variant thereof.

In a third aspect, the invention provides a purified nucleic acid molecule which encodes a polypeptide of the first or second aspects of the invention.

In one embodiment of the third aspect of the invention, the purified nucleic acid molecule preferably comprises or consists of the nucleic acid sequence as recited in SEQ ID NO:1 (encoding the INPIONCH05 exon 1 polypeptide), SEQ ID NO:3 (encoding the INPIONCH05 exon 2 polypeptide), SEQ ID NO:5 (encoding the INPIONCH05 exon 3 polypeptide), SEQ ID NO:7 (encoding the INPIONCH05 exon 4 polypeptide), SEQ ID NO:9 (encoding the INPIONCH05 exon 5 polypeptide), SEQ ID NO:11 (encoding the INPIONCH05 exon 6 polypeptide), SEQ ID NO:13 (encoding the INPIONCH05 exon 7 polypeptide), SEQ ID NO:15 (encoding the INPIONCH05 exon 8 polypeptide), SEQ ID NO:17 (encoding the INPIONCH05 exon 9 polypeptide), SEQ ID NO:19 (encoding the INPIONCH05 exon 10 polypeptide) or SEQ ID NO:21 (encoding the INPIONCH05 polypeptide), or is a redundant equivalent or fragment of any one of these sequences.

Combining the sequences recited in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17 and SEQ ID NO:19 produces the sequence recited in SEQ ID NO:21.

In a second embodiment of the third aspect of the invention, the purified nucleic acid molecule preferably comprises or consists of the nucleic acid sequence as recited in SEQ ID NO:23 (encoding the INPIONCH06 exon 1 polypeptide), SEQ ID NO:25 (encoding the INPIONCH06 exon 2 polypeptide), SEQ ID NO:27 (encoding the INPIONCH06 exon 3 polypeptide), SEQ ID NO:29 (encoding the INPIONCH06 exon 4 polypeptide), SEQ ID NO:31 (encoding the INPIONCH06 exon 5 polypeptide), SEQ ID NO:33 (encoding the

INPIONCH06 exon 6 polypeptide), SEQ ID NO:35 (encoding the INPIONCH06 exon 7 polypeptide), SEQ ID NO:37 (encoding the INPIONCH06 exon 8 polypeptide), SEQ ID NO:39 (encoding the INPIONCH06 exon 9 polypeptide), SEQ ID NO:41 (encoding the INPIONCH06 exon 10 polypeptide), SEQ ID NO:43 (encoding the INPIONCH06 exon 11 polypeptide) or SEQ ID NO:45 (encoding the INPIONCH06 polypeptide), or is a redundant equivalent or fragment of any one of these sequences.

Combining the sequences recited in SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41 and SEQ ID NO:43 produces the sequence recited in SEQ ID NO:45.

In a fourth aspect, the invention provides a purified nucleic acid molecule which hybridises under high stringency conditions with a nucleic acid molecule of the third aspect of the invention.

In a fifth aspect, the invention provides a vector, such as an expression vector, that contains a nucleic acid molecule of the third or fourth aspect of the invention.

In a sixth aspect, the invention provides a host cell transformed with a vector of the fifth aspect of the invention.

Na^+, Ca^{2+} channels and K^+ channels comprise either a four repeating ion_trans domain or a tetrameric assembly. It is considered that the CatSper channels may also adopt this conformation. In seventh aspect, the invention thus provides a tetrameric cation channel comprising a polypeptide according to the first or second aspect of the invention. The rationale for this is as follows. INPIONCH05 is covered by 2 human testis-specific ESTs. The mouse orthologue INPIONCH06 is covered by 8 mouse testis-specific ESTs. The fact that INPIONCH05 and INPIONCH06 are covered by testis-specific mRNAs and that their closest single ion_trans domain family members are the CatSper1 human and mouse sequences respectively, suggests a role for INPIONCH05 in humans and INPIONCH06 in mouse as part of a multi-subunit assembly, potentially with the other known CatSper channels or as a non-selective cationic channel involved in sperm development and/or some component of the fertilisation process. Preferably, the cation channel is a calcium-specific cation channel.

It is likely that coiled-coil motifs present in each of the four component channels act as the site of subunit interaction as they are well characterised in a number of receptors and proteins

as an oligomerisation fold, including GABAB receptors. and hemagglutinin (Kammerer R.A. *et al.* Heterodimerisation of a functional GABAB receptor is mediated by parallel coiled coil alpha helices. *Biochemistry,* 38(40): 13263-13269. (1999); Burkhard,P. *et al.* Coiled coils: a highly versatile protein folding motif. *Trends in Cell Biology.* Feb 11(2), 82-88 (2001)). A coiled-coil motif has now been identified in both the public CatSper channels and in INPIONCH05 and INPIONCH06. This coiled-coil motif is located towards the C-terminal of each of the known CatSper channel subunits. The presence of this coiled-coil motif in CatSper 1, 2 and 3 is not mentioned in the prior art.

The tetrameric assembly according to the seventh aspect of the invention could be composed of either four identical subunits, a trimer and a monomer, two pairs of homodimers, a homodimer together with two additional different subunits or finally four different subunits.

In one embodiment of the seventh aspect of the invention, the tetramer comprises either i) four INPIONCH05 monomers, ii) a trimer of INPIONCH05 and a monomer of a different polypeptide, iii) a homodimer of INPIONCH05 and a homodimer of a different polypeptide, iv) a homodimer of INPIONCH05 and two monomers of two different polypeptides, v) a monomer of INPIONCH05 and a trimer of a different polypeptide, vi) a monomer of INPIONCH05, a dimer of a different polypeptide and a monomer of a further different polypeptide, or vii) a monomer of INPIONCH05 and one monomer each of three different polypeptides, wherein the different polypeptides are preferably selected from the group consisting of human CatSper 1 (GenBank accession:AAL14105), human CatSper 2 (GenBank accession: AAL26490, AAL26491 and AAL26492) and human CatSper 3 (WO200066735-A2).

Thus a tetramer according to the first embodiment of the seventh aspect of the invention preferably comprises:

INPIONCH05-INPIONCH05-INPIONCH05-INPIONCH05;
INPIONCH05-INPIONCH05-INPIONCH05-human CatSper1;
INPIONCH05-INPIONCH05-INPIONCH05-human CatSper2;
INPIONCH05-INPIONCH05-INPIONCH05-human CatSper3;
INPIONCH05-INPIONCH05-human CatSper1-human CatSper1;
INPIONCH05-INPIONCH05-human CatSper2-human CatSper2;
INPIONCH05-INPIONCH05-human CatSper3-human CatSper3;
INPIONCH05-INPIONCH05-human CatSper1-human CatSper2;

INPIONCH05-INPIONCH05-human CatSper1-human CatSper3;
INPIONCH05-INPIONCH05-human CatSper2-human CatSper3;
INPIONCH05-human CatSper1-human CatSper1-human CatSper1;
INPIONCH05-human CatSper2-human CatSper2-human CatSper2;
INPIONCH05-human CatSper3-human CatSper3-human CatSper3;
INPIONCH05-human CatSper1-human CatSper1-human CatSper2;
INPIONCH05-human CatSper1-human CatSper1-human CatSper3;
INPIONCH05-human CatSper2-human CatSper2-human CatSper1;
INPIONCH05-human CatSper2-human CatSper2-human CatSper3;
INPIONCH05-human CatSper3-human CatSper3-human CatSper1;
INPIONCH05-human CatSper3-human CatSper3-human CatSper2; or
INPIONCH05-human CatSper1-human CatSper2-human Catsper3.

In a second embodiment of the seventh aspect of the invention, the tetramer comprises either i) four INPIONCH06 monomers, ii) a trimer of INPIONCH06 and a monomer of a different polypeptide, iii) a homodimer of INPIONCH06 and a homodimer of a different polypeptide, iv) a homodimer of INPIONCH06 and two monomers of two different polypeptides, v) a monomer of INPIONCH06, a dimer of a different polypeptide and a monomer of a further different polypeptide, or vi) a monomer of INPIONCH06 and one monomer each of three different polypeptides, wherein the different polypeptides are preferably selected from the group consisting of mouse CatSper 1 (Genbank accession: AAL14104), mouse CatSper 2 (Genbank accession: AAL26489) and mouse CatSper 3 (Genbank accession: AK014942).

Thus a tetramer according to the second embodiment of the seventh aspect of the invention preferably comprises:

INPIONCH06-INPIONCH06-INPIONCH06-INPIONCH06;
INPIONCH06-INPIONCH06-INPIONCH06-mouse CatSper1;
INPIONCH06-INPIONCH06-INPIONCH06-mouse CatSper2;
INPIONCH06-INPIONCH06-INPIONCH06-mouse CatSper3;
INPIONCH06-INPIONCH06-mouse CatSper1-mouse CatSper1;
INPIONCH06-INPIONCH06-mouse CatSper2-mouse CatSper2;
INPIONCH06-INPIONCH06-mouse CatSper3-mouse CatSper3;
INPIONCH06-INPIONCH06-mouse CatSper1-mouse CatSper2;
INPIONCH06-INPIONCH06-mouse CatSper1-mouse CatSper3;

INPIONCH06-INPIONCH06-mouse CatSper2-mouse CatSper3;
INPIONCH06-mouse CatSper1-mouse CatSper1-mouse CatSper1;
INPIONCH06-mouse CatSper2-mouse CatSper2-mouse CatSper2;
INPIONCH06-mouse CatSper3-mouse CatSper3-mouse CatSper3;
INPIONCH06-mouse CatSper1-mouse CatSper1-mouse CatSper2;
INPIONCH06-mouse CatSper1-mouse CatSper1-mouse CatSper3;
INPIONCH06-mouse CatSper2-mouse CatSper2-mouse CatSper1;
INPIONCH06-mouse CatSper2-mouse CatSper2-mouse CatSper3;
INPIONCH06-mouse CatSper3-mouse CatSper3-mouse CatSper1;
INPIONCH06-mouse CatSper3-mouse CatSper3-mouse CatSper2; or
INPIONCH06-mouse CatSper1-mouse CatSper2-mouse Catsper3.

In the tetrameric oligomers according to this aspect of the invention, any one or more of CatSper1, CatSper2, CatSper3 and a polypeptide according to first or second aspect of the invention (such as wild type INPIONCH05 or INPIONCH06 proteins whose sequences are listed herein) may be substituted by a splice variant in the channel. Splice variants have already been observed for CatSper2.

Preferably, tetrameric cation channels comprising a polypeptide according to the first aspect of the invention will be applicable to humans whilst tetrameric cation channels comprising a polypeptide according to the second aspect of the invention will be applicable to mice. It will be known to those skilled in the art that a human tetrameric channel will preferably comprise the human homologues of its respective subunits whilst a mouse tetrameric channel will preferably comprise the mouse homologues of its respective subunits. However, tetramers comprising a combination of human and mouse subunits also form part of the present invention and may be useful, for example, in research into structure/function relationships.

In an eighth aspect, the invention provides a ligand which binds specifically to, and which preferably inhibits the activity of a polypeptide of the first or second aspect of the invention or the activity of a tetrameric cation channel of the seventh aspect of the invention.

By "the activity of a polypeptide (of the invention)" or "the activity of a tetrameric cation channel (of the invention)" and similar expressions, we refer to activity characteristic of a voltage-gated cation channel and preferably activity characteristic of a member of the CatSper family of sperm-specific cation channels.

Ligands to a polypeptide according to the invention may come in various forms, including

natural or modified substrates, enzymes, receptors, small organic molecules such as small natural or synthetic organic molecules of up to 2000Da, preferably 800Da or less, peptidomimetics, inorganic molecules, peptides, polypeptides, antibodies, structural or functional mimetics of the aforementioned.

In a ninth aspect, the invention provides a compound that is effective to alter the expression of a natural gene which encodes a polypeptide of the first or second aspect of the invention or to regulate the activity of a polypeptide of the first or second aspect of the invention or a tetrameric cation channel of the seventh aspect of the invention.

A compound of the ninth aspect of the invention may either increase (agonise) or decrease (antagonise) the level of expression of the gene or the activity of the polypeptide or tetrameric cation channel.

Importantly, the identification of the function of the human INPIONCH05 and mouse INPIONCH06 polypeptides allows for the design of screening methods capable of identifying compounds that are effective in the control of voltage-gated cation channels and which are therefore effective in treating disorders associated with voltage-gated cation channels such as neurological disorders, cardiovascular disorders and autoimmune disorders. In particular, the finding that INPIONCH05 is present on chromosome 1p36.11 and that 1p36 is a known stroke locus, allows for the design of screening methods capable of identifying compounds that are effective in the control of voltage-gated cation channels which may be implicated in stroke pathology and stroke-related disorders. More importantly, the identification of the function of the human INPIONCH05 and mouse INPIONCH06 polypeptides as being similar to the function of known CatSper channel subunits allows for the design of screening methods capable of identifying compounds that are effective in controlling fertility such as in the treatment of infertility or as non-hormonal contraceptives for both men and women. Ligands and compounds according to the eighth and ninth aspects of the invention may be identified using such methods. These methods are included as aspects of the present invention.

In a tenth aspect, the invention provides a polypeptide of the first or second aspect of the invention, or a nucleic acid molecule of the third or fourth aspect of the invention, or a vector of the fifth aspect of the invention, or a ligand of the eighth aspect of the invention, or a compound of the ninth aspect of the invention, for use in therapy or diagnosis of conditions in which cation channels are implicated, such as infertility, neurological disorders, cardiovascular disorders, autoimmune disorders or stroke. These molecules may also be used

in the manufacture of a medicament for the control of fertility such as in the manufacture of fertility drugs to combat infertility or for non-hormonal contraceptives or in the manufacture of a medicament for other voltage-gated channel-related diseases and disorders such as neurological disorders, cardiovascular disorders or autoimmune disease or even stroke. These molecules may also be used in other conditions in which cation channels are implicated.

In an eleventh aspect, the invention provides a method of diagnosing a disease or disorder, including but not limited to infertility, neurological, cardiovascular, autoimmune or stroke-related disorders in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide of the first or second aspect of the invention or the activity of a polypeptide of the first or second aspect of the invention or the expression or activity or a tetrameric cationic channel according to the seventh aspect of the invention in tissue from said patient and comparing said level of expression or activity to a control level, wherein a level that is different to said control level is indicative of the disorder or the disease.

Such a method will preferably be carried out *in vitro*. Where the level of expression of a natural gene encoding a polypeptide of the first aspect of the invention or the activity or a polypeptide of the first aspect of the invention or the expression or activity of a tetrameric cationic channel according to the seventh aspect of the invention that comprises a polypeptide according to the first aspect of the invention is being assessed, this method will preferably be applied to human samples. Where the level of expression of a natural gene encoding a polypeptide of the second aspect of the invention or the activity or a polypeptide of the second aspect of the invention or the expression or activity of a tetrameric cationic channel according to the seventh aspect of the invention that comprises a polypeptide according to the second aspect of the invention is being assessed, this method will preferably be applied to mouse samples. Similar methods may be used for monitoring the therapeutic treatment of disease in a patient, wherein altering the level of expression or activity of a polypeptide or nucleic acid molecule over the period of time towards a control level is indicative of regression of the disorder or disease.

A preferred method for detecting polypeptides of the first or second aspect of the invention or tetrameric cation channels of the seventh aspect of the invention comprises the steps of: (a) contacting a ligand, such as an antibody, of the eighth aspect of the invention with a biological sample under conditions suitable for the formation of a ligand-polypeptide complex; and (b) detecting said complex.

A number of different methods according to the eleventh aspect of the invention exist, as the skilled reader will be aware, such as methods of nucleic acid hybridisation with short probes, point mutation analysis, polymerase chain reaction (PCR) amplification and methods using antibodies to detect aberrant protein levels. Similar methods may be used on a short or long term basis to allow therapeutic treatment of a disorder or disease to be monitored in a patient. The invention also provides kits that are useful in these methods for diagnosing a disorder or disease.

Preferably, the disease or disorder diagnosed by a method of the eleventh aspect of the invention is a condition in which cation channels are implicated, as described above.

In a twelfth aspect, the invention provides for the use of the polypeptides of the first or second aspects of the invention or of tetrameric cation channels of the seventh aspect of the invention as voltage-gated cation channels, preferably as sperm-specific cation channels and even more preferably as sperm-specific calcium channels, for example, as members of the CatSper family. The invention also provides for the use of a nucleic acid molecule according to the third or fourth aspects of the invention to express a protein that possesses voltage-gated cation channel activity, preferably sperm-specific cation channel activity and even more preferably sperm-specific calcium channel activity such as CatSper family activity. The invention also provides a method for effecting voltage-gated cation channel activity, sperm-specific cation channel activity and/or sperm-specific calcium channel activity, said method utilising a polypeptide of the first or second aspect of the invention.

In a thirteenth aspect, the invention provides a pharmaceutical composition comprising a polypeptide of the first or second aspect of the invention, or a nucleic acid molecule of the third or fourth aspect of the invention, or a vector of the fifth aspect of the invention, or a host cell of the sixth aspect of the invention, or a tetrameric cation channel or the seventh aspect of the invention, or a ligand of the eighth aspect of the invention, or a compound of the ninth aspect of the invention, in conjunction with a pharmaceutically-acceptable carrier.

In a fourteenth aspect, the present invention provides a polypeptide of the first or second aspect of the invention, or a nucleic acid molecule of the third or fourth aspect of the invention, or a vector of the fifth aspect of the invention, or a host cell of the sixth aspect of the invention, or a tetrameric cation channel of the seventh aspect of the invention, or a ligand of the eighth aspect of the invention, or a compound of the ninth aspect of the invention, for use in therapy of a disease or disorder, contraception or diagnosis. These molecules are

preferably used in the manufacture of a medicament for the control of fertility such as in the manufacture of fertility drugs to combat infertility or for non-hormonal contraceptives or even in the manufacture of a medicament for other voltage-gated cation channel-related diseases and disorders such as neurological disorders, cardiovascular disorders, autoimmune disease and stroke as well as other conditions in which voltage-gated cation channels, preferably sperm-specific cation channels and even more preferably sperm-specific calcium channels are implicated.

In a fifteenth aspect, the invention provides a method of contraception or a method of treating a disease or a disorder in a patient comprising administering to the patient a polypeptide of the first or second aspect of the invention, or a nucleic acid molecule of the third or fourth aspect of the invention, or a vector of the fifth aspect of the invention, or a host cell of the sixth aspect of the invention, or a tetrameric cation channel of the seventh aspect of the invention, or a ligand of the eighth aspect of the invention, or a compound of the ninth aspect of the invention.

In the case that the patient is a human, all aspects of the invention mentioned in the statement for the fifteenth aspect of the invention will be concerned with a polypeptide according to the first aspect of the invention, i.e. will be concerned with the human homologue. In the case that the subject is a mouse, all aspects of the invention mentioned in the fifteenth aspect of the invention will be concerned with a polypeptide according to the second aspect of the invention, i.e. will be concerned with the mouse homologue.

For diseases or conditions in which the expression of a natural gene encoding a polypeptide of the first or second aspect of the invention, or in which the activity of a polypeptide of the first or second aspect of the invention, or in which the expression or activity of a tetrameric cation channel according to the seventh aspect of the invention is lower in a diseased patient or a patient affected with a disorder when compared to the level of expression or activity in a healthy or non-affected patient, the polypeptide, nucleic acid molecule, vector, host cell, ligand or compound administered to the patient should be an agonist. Conversely, for diseases or conditions in which the expression of the natural gene or activity of the polypeptide or expression or activity of the tetrameric cation channel is higher in a diseased or affected patient when compared to the level of expression or activity in a healthy or non-affected patient, the polypeptide, nucleic acid molecule, vector, host cell, ligand or compound administered to the patient should be an antagonist. Examples of such antagonists include

antisense nucleic acid molecules, ribozymes and ligands, such as antibodies.

In a sixteenth aspect, the invention provides transgenic or knockout non-human animals that have been transformed to express higher, lower or absent levels of a polypeptide of the first or second aspect of the invention. Such transgenic animals are very useful models for the study of disease and disorders and may also be used in screening regimes for the identification of compounds that are effective in the treatment or diagnosis of such a disease or disorder or that are effective as contraceptive agents. Preferably, the disease or disorder is one in which cation channels are implicated, such as infertility, neurological disorders, cardiovascular disorders, autoimmune disorders or stroke.

It should be appreciated that the scope of protection sought for the polypeptides and nucleic acids of the present invention does not extend to nucleic acids or polypeptides present in their natural source. Rather, the polypeptides and nucleic acids claimed by the present invention may be regarded as being "isolated" or "purified". The terms "isolated" and "purified" as used herein refer to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. Thus, for example, a polypeptide contained in a tissue extract would constitute an "isolated" or "purified" polypeptide, as would a polypeptide synthetically or recombinantly produced. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same.

It should be noted that the terms "isolated" and "purified" do not denote the method by which the polypeptide or nucleic acid is obtained or the level of purity of the preparation. Thus, such isolated or purified species may be produced recombinantly, isolated directly from the cell or tissue of interest or produced synthetically based on the determined sequences.

A summary of standard techniques and procedures which may be employed in order to utilise the invention is given below. It will be understood that this invention is not limited to the particular methodology, protocols, cell lines, vectors and reagents described. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and it is not intended that this terminology should limit the scope of the present invention. The extent of the invention is limited only by the terms of the appended claims.

Standard abbreviations for nucleotides and amino acids are used in this specification.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA technology and immunology, which are within the skill of those working in the art.

Such techniques are explained fully in the literature. Examples of particularly suitable texts for consultation include the following: Sambrook Molecular Cloning; A Laboratory Manual, Second Edition (1989); DNA Cloning, Volumes I and II (D.N Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames & S.J. Higgins eds. 1984); Animal Cell Culture (R.I. Freshney ed. 1986); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide to Molecular Cloning (1984); the Methods in Enzymology series (Academic Press, Inc.), especially volumes 154 & 155; Gene Transfer Vectors for Mammalian Cells (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Immunochemical Methods in Cell and Molecular Biology (Mayer and Walker, eds. 1987, Academic Press, London); Scopes, (1987) Protein Purification: Principles and Practice, Second Edition (Springer Verlag, N.Y.); and Handbook of Experimental Immunology, Volumes I-IV (D.M. Weir and C. C. Blackwell eds. 1986).

As used herein, the term "polypeptide" includes any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e. peptide isosteres. This term refers both to short chains (peptides and oligopeptides) and to longer chains (proteins).

The polypeptide of the present invention may be in the form of a mature protein or may be a pre-, pro- or prepro- protein that can be activated by cleavage of the pre-, pro- or prepro- portion to produce an active mature polypeptide. In such polypeptides, the pre-, pro- or prepro- sequence may be a leader or secretory sequence or may be a sequence that is employed for purification of the mature polypeptide sequence.

The polypeptide of the first or second aspect of the invention may form part of a fusion protein. For example, it is often advantageous to include one or more additional amino acid sequences which may contain secretory or leader sequences, pro-sequences, sequences which aid in purification, or sequences that confer higher protein stability, for example during recombinant production. Alternatively or additionally, the mature polypeptide may be fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol).

Polypeptides may contain amino acids other than the 20 gene-encoded amino acids, modified either by natural processes, such as by post-translational processing or by chemical modification techniques which are well known in the art. Among the known modifications which may commonly be present in polypeptides of the present invention are glycosylation, lipid attachment, sulphation, gamma-carboxylation, for instance of glutamic acid residues, hydroxylation and ADP-ribosylation. Other potential modifications include acetylation, acylation, amidation, covalent attachment of flavin, covalent attachment of a haeme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulphide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, GPI anchor formation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl terminus in a polypeptide, or both, by a covalent modification is common in naturally-occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention.

The modifications that occur in a polypeptide often will be a function of how the polypeptide is made. For polypeptides that are made recombinantly, the nature and extent of the modifications in large part will be determined by the post-translational modification capacity of the particular host cell and the modification signals that are present in the amino acid sequence of the polypeptide in question. For instance, glycosylation patterns vary between different types of host cell.

The polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally-occurring polypeptides (for example purified from cell culture), recombinantly-produced polypeptides (including fusion proteins), synthetically-produced polypeptides or polypeptides that are produced by a combination of these methods.

The functionally-equivalent polypeptides of the first or second aspect of the invention may be polypeptides that are homologous to the INPIONCH05 and INPIONCH06 polypeptides. Two polypeptides are said to be "homologous", as the term is used herein, if the sequence of one of

the polypeptides has a high enough degree of identity or similarity to the sequence of the other polypeptide. "Identity" indicates that at any particular position in the aligned sequences, the amino acid residue is identical between the sequences. "Similarity" indicates that, at any particular position in the aligned sequences, the amino acid residue is of a similar type between the sequences. Degrees of identity and similarity can be readily calculated (Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing. Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part 1, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991).

Homologous polypeptides therefore include natural biological variants (for example, allelic variants or geographical variations within the species from which the polypeptides are derived) and mutants (such as mutants containing amino acid substitutions, insertions or deletions) of the INPIONCH05 and INPIONCH06 polypeptides. Such mutants may include polypeptides in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; among the basic residues Lys and Arg; or among the aromatic residues Phe and Tyr. Particularly preferred are variants in which several, i.e. between 5 and 10, 1 and 5, 1 and 3, 1 and 2 or just 1 amino acids are substituted, deleted or added in any combination. Especially preferred are silent substitutions, additions and deletions, which do not alter the properties and activities of the protein. Also especially preferred in this regard are conservative substitutions. Such mutants also include polypeptides in which one or more of the amino acid residues includes a substituent group.

Typically, greater than 30% identity between two polypeptides is considered to be an indication of functional equivalence. Preferably, functionally equivalent polypeptides of the first or second aspect of the invention have a degree of sequence identity with the INPIONCH05 and INPIONCH06 polypeptides, or with active fragments thereof, of greater than 80%. More preferred polypeptides have degrees of identity of greater than 85%, 90%, 95%, 98% or 99%, respectively.

The functionally-equivalent polypeptides of the first or second aspect of the invention may also be polypeptides which have been identified using one or more techniques of structural alignment. For example, the Inpharmatica Genome Threader™ technology that forms one aspect of the search tools used to generate the Biopendium search database may be used (see co-pending International Patent Application No. PCT/GB01/01105) to identify polypeptides of presently-unknown function which, while having low sequence identity as compared to the INPIONCH05 and INPIONCH06 polypeptides, are predicted to be voltage-gated cation channels, preferably sperm-specific cation channels and even more preferably sperm-specific calcium channels, said method utilising a polypeptide of the first or second aspect of the invention, by virtue of sharing significant structural homology with the INPIONCH05 and INPIONCH06 polypeptide sequences. By "significant structural homology" is meant that the Inpharmatica Genome Threader™ predicts two proteins to share structural homology with a certainty of at least 10% and more preferably, at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% and above.

The polypeptides of the first and second aspects of the invention also include fragments of the INPIONCH05 and INPIONCH06 polypeptides and fragments of the functional equivalents of the INPIONCH05 and INPIONCH06 polypeptides, provided that those fragments retain voltage-gated cation channel activity, preferably sperm-specific voltage-gated cation channel activity and more preferably sperm-specific calcium channel activity such as that of the CatSper family or have an antigenic determinant in common with the INPIONCH05 and INPIONCH06 polypeptides.

As used herein, the term "fragment" refers to a polypeptide having an amino acid sequence that is the same as part, but not all, of the amino acid sequence of the INPIONCH05 and INPIONCH06 polypeptides or one of its functional equivalents. The fragments should comprise at least n consecutive amino acids from the sequence and, depending on the particular sequence, n preferably is 7 or more (for example, 8, 10, 12, 14, 16, 18, 20 or more). Small fragments may form an antigenic determinant.

Fragments of the full length INPIONCH05 and INPIONCH06 polypeptides may consist of combinations of 1, 2, 3, 4, 5 or more neighbouring exon sequences in the INPIONCH05 and INPIONCH06 polypeptide sequences, respectively. For example, such combinations include exons 1 and 2, exons 2 and 3, or exons 2 to 4, and so on. Such fragments are included in the present invention.

Such fragments may be "free-standing", i.e. not part of or fused to other amino acids or polypeptides, or they may be comprised within a larger polypeptide of which they form a part or region. When comprised within a larger polypeptide, the fragment of the invention most preferably forms a single continuous region. For instance, certain preferred embodiments relate to a fragment having a pre - and/or pro- polypeptide region fused to the amino terminus of the fragment and/or an additional region fused to the carboxyl terminus of the fragment. However, several fragments may be comprised within a single larger polypeptide.

The polypeptides of the present invention or their immunogenic fragments (comprising at least one antigenic determinant) or the tetrameric cation channels of the seventh aspect of the invention can be used to generate ligands, such as polyclonal or monoclonal antibodies, that are immunospecific for the polypeptides or the tetrameric ion channels respectively. Such antibodies may be employed to isolate or to identify clones expressing the polypeptides of the invention or to purify the polypeptides of the tetrameric cation channels by affinity chromatography. The antibodies may also be employed as diagnostic or therapeutic aids, amongst other applications, as will be apparent to the skilled reader.

The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides or the tetrameric cation channels of the invention than their affinity for other related polypeptides in the prior art. As used herein, the term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$ and Fv, which are capable of binding to the antigenic determinant in question. Such antibodies thus bind to the polypeptides of the first and second aspects of the invention and/or the tetrameric cation channels of the seventh aspect of the invention.

By "substantially greater affinity" we mean that there is a measurable increase in the affinity for a polypeptide of the invention as compared with the affinity for known cell-surface receptors.

Preferably, the affinity is at least 1.5-fold, 2-fold, 5-fold 10-fold, 100-fold, 10^3-fold, 10^4-fold, 10^5-fold or 10^6-fold greater for a polypeptide of the invention than for known cell-surface receptor polypeptides.

If polyclonal antibodies are desired, a selected mammal, such as a mouse (not for INPIONCH06), rabbit, goat or horse, may be immunised with a polypeptide of the first or second aspects of the invention or a tetrameric cation channel of the seventh aspect of the invention. The polypeptide used to immunise the animal can be derived by recombinant DNA

technology or can be synthesised chemically. If desired, the polypeptide can be conjugated to a carrier protein. Commonly used carriers to which the polypeptides may be chemically coupled include bovine serum albumin, thyroglobulin and keyhole limpet haemocyanin. The coupled polypeptide is then used to immunise the animal. Serum from the immunised animal is collected and treated according to known procedures, for example by immunoaffinity chromatography.

Monoclonal antibodies to the polypeptides of the first or second aspects of the invention or to the tetrameric cation channels of the seventh aspect of the invention can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies using hybridoma technology is well known (see, for example, Kohler, G. and Milstein, C., Nature 256: 495-497 (1975); Kozbor *et al.*, Immunology Today 4: 72 (1983); Cole *et al.*, 77-96 in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc. (1985).

Panels of monoclonal antibodies produced against the polypeptides of the first aspect and/or second aspects of the invention or the tetrameric cation channels of the seventh aspect of the invention can be screened for various properties, i.e., for isotype, epitope, affinity, etc. Monoclonal antibodies are particularly useful in purification of the individual polypeptides or tetrameric cation channels against which they are directed. Alternatively, genes encoding the monoclonal antibodies of interest may be isolated from hybridomas, for instance by PCR techniques known in the art, and cloned and expressed in appropriate vectors.

Chimeric antibodies, in which non-human variable regions are joined or fused to human constant regions (see, for example, Liu *et al.*, Proc. Natl. Acad. Sci. USA, 84, 3439 (1987)), may also be of use.

The antibody may be modified to make it less immunogenic in an individual, for example by humanisation (see Jones *et al.*, Nature, 321, 522 (1986); Verhoeyen *et al.*, Science, 239, 1534 (1988); Kabat *et al.*, J. Immunol., 147, 1709 (1991); Queen *et al.*, Proc. Natl Acad. Sci. USA, 86, 10029 (1989); Gorman *et al.*, Proc. Natl Acad. Sci. USA, 88, 34181 (1991); and Hodgson *et al.*, Bio/Technology, 9, 421 (1991)). The term "humanised antibody", as used herein, refers to antibody molecules in which the CDR amino acids and selected other amino acids in the variable domains of the heavy and/or light chains of a non-human donor antibody have been substituted in place of the equivalent amino acids in a human antibody. The humanised antibody thus closely resembles a human antibody but has the binding ability of the donor antibody.

In a further alternative, the antibody may be a "bispecific" antibody, that is an antibody having two different antigen binding domains, each domain being directed against a different epitope.

Phage display technology may be utilised to select genes which encode antibodies with binding activities towards the polypeptides of the invention either from repertoires of PCR amplified V-genes of lymphocytes from humans screened for possessing the relevant antibodies, or from naive libraries (McCafferty, J. *et al.*, (1990), Nature 348, 552-554; Marks, J. *et al.*, (1992) Biotechnology 10, 779-783). The affinity of these antibodies can also be improved by chain shuffling (Clackson, T. *et al.*, (1991) Nature 352, 624-628).

Antibodies generated by the above techniques, whether polyclonal or monoclonal, have additional utility in that they may be employed as reagents in immunoassays, radioimmunoassays (RIA) or enzyme-linked immunosorbent assays (ELISA). In these applications, the antibodies can be labelled with an analytically-detectable reagent such as a radioisotope, a fluorescent molecule or an enzyme.

Preferred nucleic acid molecules of the third and fourth aspects of the invention are those which encode the polypeptide sequences recited in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20 and/or SEQ ID NO:22, or SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44 and/or SEQ ID NO:46, and functionally equivalent polypeptides. These nucleic acid molecules may be used in the methods and applications described herein. The nucleic acid molecules of the invention preferably comprise at least n consecutive nucleotides from the sequences disclosed herein where, depending on the particular sequence, n is 10 or more (for example, 12, 14, 15, 18, 20, 25, 30, 35, 40 or more).

The nucleic acid molecules of the invention also include sequences that are complementary to nucleic acid molecules described above (for example, for antisense or probing purposes).

Nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance cDNA, synthetic DNA or genomic DNA. Such nucleic acid molecules may be obtained by cloning, by chemical synthetic techniques or by a combination thereof. The nucleic acid molecules can be prepared, for example, by chemical synthesis using techniques such as solid phase phosphoramidite chemical synthesis, from

genomic or cDNA libraries or by separation from an organism. RNA molecules may generally be generated by the *in vitro* or *in vivo* transcription of DNA sequences.

The nucleic acid molecules may be double-stranded or single-stranded. Single-stranded DNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

The term "nucleic acid molecule" also includes analogues of DNA and RNA, such as those containing modified backbones, and peptide nucleic acids (PNA). The term "PNA", as used herein, refers to an antisense molecule or an anti-gene agent which comprises an oligonucleotide of at least five nucleotides in length linked to a peptide backbone of amino acid residues, which preferably ends in lysine. The terminal lysine confers solubility to the composition. PNAs may be pegylated to extend their lifespan in a cell, where they preferentially bind complementary single stranded DNA and RNA and stop transcript elongation (Nielsen, P.E. *et al.* (1993) Anticancer Drug Des. 8:53-63).

These molecules also may have a different sequence which, as a result of the degeneracy of the genetic code, encode a polypeptide of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20 and/or SEQ ID NO:22, or SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44 and/or SEQ ID NO:46. Such molecules may include, but are not limited to, the coding sequence for the mature polypeptide by itself; the coding sequence for the mature polypeptide and additional coding sequences, such as those encoding a leader or secretory sequence, such as a pro-, pre- or prepro- polypeptide sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences, together with further additional, non-coding sequences, including non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription (including termination signals), ribosome binding and mRNA stability. The nucleic acid molecules may also include additional sequences which encode additional amino acids, such as those which provide additional functionalities.

The nucleic acid molecules of the third and fourth aspects of the invention may also encode the fragments or the functional equivalents of the polypeptides and fragments of the first and second aspects of the invention. Such a nucleic acid molecule may be a naturally-occurring variant such as a naturally-occurring allelic variant, or the molecule may be a variant that is

not known to occur naturally. Such non-naturally occurring variants of the nucleic acid molecule may be made by mutagenesis techniques, including those applied to nucleic acid molecules, cells or organisms.

Among variants in this regard are variants that differ from the aforementioned nucleic acid molecules by nucleotide substitutions, deletions or insertions. The substitutions, deletions or insertions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or insertions.

The nucleic acid molecules of the invention can also be engineered, using methods generally known in the art, for a variety of reasons, including modifying the cloning, processing, and/or expression of the gene product (the polypeptide). DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides are included as techniques which may be used to engineer the nucleotide sequences. Site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations and so forth.

Nucleic acid molecules which encode a polypeptide of the first or second aspect of the invention may be ligated to a heterologous sequence so that the combined nucleic acid molecule encodes a fusion protein. Such combined nucleic acid molecules are included within the third or fourth aspects of the invention. For example, to screen peptide libraries for inhibitors of the activity of the polypeptide or of the tetrameric cation channel of the seventh aspect of the invention, it may be useful to express, using such a combined nucleic acid molecule, a fusion protein that can be recognised by a commercially-available antibody. A fusion protein may also be engineered to contain a cleavage site located between the sequence of the polypeptide of the invention and the sequence of a heterologous protein so that the polypeptide may be cleaved and purified away from the heterologous protein.

The nucleic acid molecules of the invention also include antisense molecules that are partially complementary to nucleic acid molecules encoding polypeptides of the present invention and that therefore hybridise to the encoding nucleic acid molecules (hybridisation). Such antisense molecules, such as oligonucleotides, can be designed to recognise, specifically bind to and prevent transcription of a target nucleic acid encoding a polypeptide of the invention, as will be known by those of ordinary skill in the art (see, for example, Cohen, J.S., Trends in Pharm. Sci., 10, 435 (1989), Okano, J. Neurochem. 56, 560 (1991); O'Connor, J. Neurochem 56, 560

(1991); Lee *et al.*, Nucleic Acids Res 6, 3073 (1979); Cooney *et al.*, Science 241, 456 (1988); Dervan *et al.*, Science 251, 1360 (1991).

The term "hybridisation" as used herein refers to the association of two nucleic acid molecules with one another by hydrogen bonding. Typically, one molecule will be fixed to a solid support and the other will be free in solution. Then, the two molecules may be placed in contact with one another under conditions that favour hydrogen bonding. Factors that affect this bonding include: the type and volume of solvent; reaction temperature; time of hybridisation; agitation; agents to block the non-specific attachment of the liquid phase molecule to the solid support (Denhardt's reagent or BLOTTO); the concentration of the molecules; use of compounds to increase the rate of association of molecules (dextran sulphate or polyethylene glycol); and the stringency of the washing conditions following hybridisation (see Sambrook *et al.* [*supra*]).

The inhibition of hybridisation of a completely complementary molecule to a target molecule may be examined using a hybridisation assay, as known in the art (see, for example, Sambrook *et al* [*supra*]). A substantially homologous molecule will then compete for and inhibit the binding of a completely homologous molecule to the target molecule under various conditions of stringency, as taught in Wahl, G.M. and S.L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A.R. (1987; Methods Enzymol. 152:507-511).

"Stringency" refers to conditions in a hybridisation reaction that favour the association of very similar molecules over association of molecules that differ. High stringency hybridisation conditions are defined as overnight incubation at 42°C in a solution comprising 50% formamide, 5XSSC (150mM NaCl, 15mM trisodium citrate), 50mM sodium phosphate (pH7.6), 5x Denhardts solution, 10% dextran sulphate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed b y washing the filters in 0.1X SSC at approximately 65°C. Low stringency conditions involve the hybridisation reaction being carried out at 35°C (see Sambrook *et al.* [*supra*]). Preferably, the conditions used for hybridisation are those of high stringency.

Preferred embodiments of this aspect of the invention are nucleic acid molecules that are at least 60% identical, preferably at least 70% identical, even more preferably at least 80% identical over their entire length to a nucleic acid molecule encoding the INPIONCH05 or INPIONCH06 polypeptides (SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID

NO:19 and/or SEQ ID NO:21, or SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43 and/or SEQ ID NO:45) and nucleic acid molecules that are substantially complementary to such nucleic acid molecules. Preferably, a nucleic acid molecule according to one embodiment of this aspect of the invention comprises a region that is at least 70%, preferably at least 80% identical over its entire length to the coding sequence for SEQ ID NO:2 given in SEQ ID NO:1, the coding sequence for SEQ ID NO:4 given in SEQ ID NO:3, the coding sequence for SEQ ID NO:6 given in SEQ ID NO:5, the coding sequence for SEQ ID NO:8 given in SEQ ID NO:7, the coding sequence for SEQ ID NO:10 given in SEQ ID NO:9, the coding sequence for SEQ ID NO:12 given in SEQ ID NO:11, the coding sequence for SEQ ID NO:14 given in SEQ ID NO:13, the coding sequence for SEQ ID NO:16 given in SEQ ID NO:15, the coding sequence for SEQ ID NO:18 given in SEQ ID NO:17, the coding sequence for SEQ ID NO:20 given in SEQ ID NO:19 or the coding sequence for SEQ ID NO:22 given in SEQ ID NO:21, or is a nucleic acid molecule that is complementary thereto. Preferably, a nucleic acid molecule according to a second embodiment of this aspect of the invention comprises a region that is at least 70% identical, preferably at least 80% identical over its entire length to the coding sequence for SEQ ID NO:24 given in SEQ ID NO:23, the coding sequence for SEQ ID NO:26 given in SEQ ID NO:25, the coding sequence for SEQ ID NO:28 given in SEQ ID NO:27, the coding sequence for SEQ ID NO:30 given in SEQ ID NO:29, the coding sequence for SEQ ID NO:32 given in SEQ ID NO:31, the coding sequence for SEQ ID NO:34 given in SEQ ID NO:33, the coding sequence for SEQ ID NO:36 given in SEQ ID NO:35, the coding sequence for SEQ ID NO:38 given in SEQ ID NO:37, the coding sequence for SEQ ID NO:40 given in SEQ ID NO:39, the coding sequence for SEQ ID NO:42 given in SEQ ID NO:41, the coding sequence for SEQ ID NO:44 given in SEQ ID NO:43 or the coding sequence for SEQ ID NO:46 given in SEQ ID NO:45 or is a nucleic acid molecule that is complementary thereto. In these two embodiments, nucleic acid molecules at least 90%, preferably at least 95%, more preferably at least 98% or 99% identical over their entire length to the same are even more particularly preferred. Preferred embodiments in this respect are nucleic acid molecules that encode polypeptides which retain substantially the same biological function or activity as the INPIONCH05 and INPIONCH06 polypeptides.

The invention also provides a process for detecting a nucleic acid molecule of the invention, comprising the steps of: (a) contacting a nucleic probe according to the invention with a

biological sample under hybridising conditions to form duplexes; and (b) detecting any such duplexes that are formed.

As discussed additionally below in connection with assays that may be utilised according to the invention, a nucleic acid molecule as described above may be used as a hybridisation probe for RNA, cDNA or genomic DNA, in order to isolate full-length cDNAs and genomic clones encoding the INPIONCH05 and INPIONCH06 polypeptides and to isolate cDNA and genomic clones of homologous or orthologous genes that have a high sequence similarity to the gene encoding these polypeptides.

In this regard, the following techniques, among others known in the art, may be utilised and are discussed below for purposes of illustration. Methods for DNA sequencing and analysis are well known and are generally available in the art and may, indeed, be used to practice many of the embodiments of the invention discussed herein. Such methods may employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase (US Biochemical Corp, Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham, Chicago, IL), or combinations of polymerases and proof-reading exonucleases such as those found in the ELONGASE Amplification System marketed by Gibco/BRL (Gaithersburg, MD). Preferably, the sequencing process may be automated using machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, NV), the Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA) and the ABI Catalyst and 373 and 377 DNA Sequencers (Perkin Elmer).

One method for isolating a nucleic acid molecule encoding a polypeptide with an equivalent function to that of the INPIONCH05 and INPIONCH06 polypeptides is to probe a genomic or cDNA library with a natural or artificially-designed probe using standard procedures that are recognised in the art (see, for example, "Current Protocols in Molecular Biology", Ausubel *et al.* (eds). Greene Publishing Association and John Wiley Interscience, New York, 1989,1992). Probes comprising at least 15, preferably at least 30, and more preferably at least 50, contiguous bases that correspond to, or are complementary to, nucleic acid sequences from the appropriate encoding gene (SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19 and/or SEQ ID NO:21, or SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43 and/or SEQ ID NO:45) are particularly useful probes. Such

probes may be labelled with an analytically-detectable reagent to facilitate their identification. Useful reagents include, but are not limited to, radioisotopes, fluorescent dyes and enzymes that are capable of catalysing the formation of a detectable product. Using these probes, the ordinarily skilled artisan will be capable of isolating complementary copies of genomic DNA, cDNA or RNA polynucleotides encoding proteins of interest from human, mammalian or other animal sources and screening such sources for related sequences, for example, for additional members of the family, type and/or subtype.

In many cases, isolated cDNA sequences will be incomplete, in that the region encoding the polypeptide will be cut short, normally at the 5' end. Several methods are available to obtain full length cDNAs, or to extend short cDNAs. Such sequences may be extended utilising a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed is based on the method of Rapid Amplification of cDNA Ends (RACE; see, for example, Frohman *et al.*, PNAS USA 85, 8998-9002, 1988). Recent modifications of this technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.), for example, have significantly simplified the search for longer cDNAs. A slightly different technique, termed "restriction-site" PCR, uses universal primers to retrieve unknown nucleic acid sequence adjacent a known locus (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). Inverse PCR may also be used to amplify or to extend sequences using divergent primers based on a known region (Triglia, T. *et al.* (1988) Nucleic Acids Res. 16:8186). Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent a known sequence in human and yeast artificial chromosome DNA (Lagerstrom, M. *et al.* (1991) PCR Methods Applic., 1, 111-119). Another method which may be used to retrieve unknown sequences is that of Parker, J.D. *et al.* (1991); Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PromoterFinderTM libraries to walk genomic DNA (Clontech, Palo Alto, CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they will contain more sequences that contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

In one embodiment of the invention, the nucleic acid molecules of the present invention may be used for chromosome localisation. In this technique, a nucleic acid molecule is specifically targeted to, and can hybridise with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important step in the confirmatory correlation of those sequences with the gene-associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationships between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes). This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localised by genetic linkage to a particular genomic region, any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleic acid molecule may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or affected individuals.

The nucleic acid molecules of the present invention are also valuable for tissue localisation. Such techniques allow the determination of expression patterns of the polypeptide in tissues by detection of the mRNAs that encode them. These techniques include *in situ* hybridisation techniques and nucleotide amplification techniques, such as PCR. Results from these studies provide an indication of the normal functions of the polypeptide in the organism. In addition, comparative studies of the normal expression pattern of mRNAs with that of mRNAs encoded by a mutant gene provide valuable insights into the role of mutant polypeptides in disease. Such inappropriate expression may be of a temporal, spatial or quantitative nature.

Gene silencing approaches may also be undertaken to down-regulate endogenous expression of a gene encoding a polypeptide of the invention. RNA interference (RNAi) (Elbashir, SM *et al.*, Nature 2001, 411, 494-498) is one method of sequence specific post-transcriptional gene silencing that may be employed. Short dsRNA oligonucleotides are synthesised *in vitro* and introduced into a cell. The sequence specific binding of these dsRNA oligonucleotides triggers the degradation of target mRNA, reducing or ablating target protein expression.

Efficacy of the gene silencing approaches assessed above may be assessed through the

measurement of polypeptide expression (for example, by Western blotting), and at the RNA level using TaqMan-based methodologies.

The vectors of the present invention comprise nucleic acid molecules of the invention and may be cloning or expression vectors. The host cells of the invention, which may be transformed, transfected or transduced with the vectors of the invention may be prokaryotic or eukaryotic.

The polypeptides of the invention may be prepared in recombinant form by expression of their encoding nucleic acid molecules in vectors contained within a host cell. Such expression methods are well known to those of skill in the art and many are described in detail by Sambrook *et al* (*supra*) and Fernandez & Hoeffler (1998, eds. "Gene expression systems. Using nature for the art of expression". Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto).

Generally, any system or vector that is suitable to maintain, propagate or express nucleic acid molecules to produce a polypeptide in the required host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those described in Sambrook *et al.*, (*supra*). Generally, the encoding gene can be placed under the control of a control element such as a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator, so that the DNA sequence encoding the desired polypeptide is transcribed into RNA in the transformed host cell.

Examples of suitable expression systems include, for example, chromosomal, episomal and virus-derived systems, including, for example, vectors derived from: bacterial plasmids, bacteriophage, transposons, yeast episomes, insertion elements, yeast chromosomal elements, viruses such as baculoviruses, papova viruses such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, or combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, including cosmids and phagemids. Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained and expressed in a plasmid.

Particularly suitable expression systems include microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (for example, baculovirus); plant cell systems transformed with virus expression

vectors (for example, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (for example, Ti or pBR322 plasmids); or animal cell systems. Cell-free translation systems can also be employed to produce the polypeptides of the invention.

Introduction of nucleic acid molecules encoding a polypeptide of the present invention into host cells can be effected by methods described in many standard laboratory manuals, such as Davis *et al.*, Basic Methods in Molecular Biology (1986) and Sambrook *et al.*,[*supra*]. Particularly suitable methods include calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection (see Sambrook *et al.*, 1989 [*supra*]; Ausubel *et al.*, 1991 [*supra*]; Spector, Goldman & Leinwald, 1998). In eukaryotic cells, expression systems may either be transient (for example, episomal) or permanent (chromosomal integration) according to the needs of the system.

The encoding nucleic acid molecule may or may not include a sequence encoding a control sequence, such as a signal peptide or leader sequence, as desired, for example, for secretion of the translated polypeptide into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment. These signals may be endogenous to the polypeptide or they may be heterologous signals. Leader sequences can be removed by the bacterial host in post-translational processing.

In addition to control sequences, it may be desirable to add regulatory sequences that allow for regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory sequences are those which cause the expression of a gene to be increased or decreased in response to a chemical or physical stimulus, including the presence of a regulatory compound or to various temperature or metabolic conditions. Regulatory sequences are those non-translated regions of the vector, such as enhancers, promoters and 5' and 3' untranslated regions. These interact with host cellular proteins to carry out transcription and translation. Such regulatory sequences may vary in their strength and specificity. Depending on the vector system and host utilised, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript phagemid (Stratagene, LaJolla, CA) or pSportl™ plasmid (Gibco BRL) and the like may be used. The baculovirus polyhedrin promoter may be used in insect

cells. Promoters or enhancers derived from the genomes of plant cells (for example, heat shock, RUBISCO and storage protein genes) or from plant viruses (for example, viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of the sequence, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

An expression vector is constructed so that the particular nucleic acid coding sequence is located in the vector with the appropriate regulatory sequences, the positioning and orientation of the coding sequence with respect to the regulatory sequences being such that the coding sequence is transcribed under the "control" of the regulatory sequences, i.e., RNA polymerase which binds to the DNA molecule at the control sequences transcribes the coding sequence. In some cases it may be necessary to modify the sequence so that it may be attached to the control sequences with the appropriate orientation; i.e., to maintain the reading frame.

The control sequences and other regulatory sequences may be ligated to the nucleic acid coding sequence prior to insertion into a vector. Alternatively, the coding sequence can be cloned directly into an expression vector that already contains the control sequences and an appropriate restriction site.

For long-term, high-yield production of a recombinant polypeptide, stable expression is preferred. For example, cell lines which stably express the polypeptide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells that successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalised cell lines available from the American Type Culture Collection (ATCC) including, but not limited to, Chinese hamster ovary (CHO), HeLa, baby hamster kidney (BHK), monkey kidney (COS), C127, 3T3, BHK, HEK 293, Bowes melanoma and human

hepatocellular carcinoma (for example Hep G2) cells and a number of other cell lines.

In the baculovirus system, the materials for baculovirus/insect cell expression systems are commercially available in kit form from, inter alia, Invitrogen, San Diego CA (the "MaxBac" kit). These techniques are generally known to those skilled in the art and are described fully in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987). Particularly suitable host cells for use in this system include insect cells such as Drosophila S2 and Spodoptera Sf9 cells.

There are many plant cell culture and whole plant genetic expression systems known in the art. Examples of suitable plant cellular genetic expression systems include those described in US 5,693,506; US 5,659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture has been described by Zenk, Phytochemistry 30, 3861-3863 (1991).

In particular, all plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be utilised, so that whole plants are recovered which contain the transferred gene. Practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugar cane, sugar beet, cotton, fruit and other trees, legumes and vegetables.

Examples of particularly preferred bacterial host cells include *Streptococci, Staphylococci, E. coli, Streptomyces* and *Bacillus subtilis* cells.

Examples of particularly suitable host cells for fungal expression include yeast cells (for example, *S. cerevisiae*) and *Aspergillus* cells.

Any number of selection systems are known in the art that may be used to recover transformed cell lines. Examples include the herpes simplex virus thymidine kinase (Wigler, M. *et al.* (1977) Cell 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. *et al.* (1980) Cell 22:817-23) genes that can be employed in tk- or aprt± cells, respectively.

Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dihydrofolate reductase (DHFR) that confers resistance to methotrexate (Wigler, M. *et al.* (1980) Proc. Natl. Acad. Sci. 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. *et al* (1981) J. Mol. Biol. 150:1-14) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. Additional selectable genes have been described, examples of which will be clear to those of skill in the art.

Although the presence or absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the relevant sequence is inserted within a marker gene sequence, transformed cells containing the appropriate sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a polypeptide of the invention under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells that contain a nucleic acid sequence encoding a polypeptide of the invention and which express said polypeptide may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridisations and protein bioassays, for example, fluorescence activated cell sorting (FACS) or immunoassay techniques (such as the enzyme-linked immunosorbent assay [ELISA] and radioimmunoassay [RIA]), that include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein (see Hampton, R. *et al.* (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN) and Maddox, D.E. *et al.* (1983) J. Exp. Med, 158, 1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labelled hybridisation or PCR probes for detecting sequences related to nucleic acid molecules encoding polypeptides of the present invention include oligolabelling, nick translation, end-labelling or PCR amplification using a labelled polynucleotide. Alternatively, the sequences encoding the polypeptide of the invention may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesise RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3 or SP6 and labelled nucleotides. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, MI); Promega (Madison WI); and U.S. Biochemical Corp., Cleveland, OH)).

Suitable reporter molecules or labels, which may be used for ease of detection, include radionuclides, enzymes and fluorescent, chemiluminescent or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Nucleic acid molecules according to the present invention may also be used to create transgenic animals, particularly rodent animals. Such transgenic animals form a further aspect

of the present invention. This may be done locally by modification of somatic cells, or by germ line therapy to incorporate heritable modifications. Such transgenic animals may be particularly useful in the generation of animal models for drug molecules effective as modulators of the polypeptides or the tetrameric cation channels of the present invention.

The polypeptide can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulphate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. High performance liquid chromatography is particularly useful for purification. Well known techniques for refolding proteins may be employed to regenerate an active conformation when the polypeptide is denatured during isolation and or purification.

Specialised vector constructions may also be used to facilitate purification of proteins, as desired, by joining sequences encoding the polypeptides of the invention to a nucleotide sequence encoding a polypeptide domain that will facilitate purification of soluble proteins. Examples of such purification-facilitating domains include metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilised metals, protein A domains that allow purification on immobilised immunoglobulin, and the domain utilised in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, WA). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the polypeptide of the invention may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing the polypeptide of the invention fused to several histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilised metal ion affinity chromatography as described in Porath, J. *et al.* (1992), Prot. Exp. Purif. 3: 263-281) while the thioredoxin or enterokinase cleavage site provides a means for purifying the polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D.J. *et al.* (1993; DNA Cell Biol. 12:441-453).

If the polypeptide is to be expressed for use in screening assays, generally it is preferred that it be produced at the surface of the host cell in which it is expressed. In this event, the host cells may be harvested prior to use in the screening assay, for example using techniques such as fluorescence activated cell sorting (FACS) or immunoaffinity techniques. If the polypeptide is

secreted into the medium, the medium can be recovered in order to recover and purify the expressed polypeptide. If polypeptide is produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

The polypeptide of the invention can be used to screen libraries of compounds in any of a variety of drug screening techniques. Such compounds may activate (agonise) or inhibit (antagonise) the level of expression of the gene or the activity of the polypeptide or the expression or activity of the tetrameric cation channel of the invention and form a further aspect of the present invention. Preferred compounds are effective to alter the expression of a natural gene which encodes a polypeptide of the first or second aspect of the invention or to regulate the activity of a polypeptide of the first or second aspect of the invention or to alter the expression of regulate the activity of a tetrameric cation channel of the seventh aspect of the invention.

Agonist or antagonist compounds may be isolated from, for example, cells, cell-free preparations, chemical libraries or natural product mixtures. These agonists or antagonists may be natural or modified substrates, ligands, enzymes, receptors or structural or functional mimetics. For a suitable review of such screening techniques, see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991).

Compounds that are most likely to be good antagonists are molecules that bind to the polypeptide of the invention or the tetrameric cation channel of the invention without inducing the biological effects of the polypeptide or the tetrameric cation channel upon binding to it. Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to the polypeptide or tetrameric cation channel of the invention and thereby inhibit or extinguish its activity. In this fashion, binding of the polypeptide or tetrameric cation channel to normal cellular binding molecules may be inhibited, such that the normal biological activity of the polypeptide or tetrameric cation channel is prevented.

The polypeptide of the invention that is employed in such a screening technique may be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. In general, such screening procedures may involve using appropriate cells or cell membranes that express the polypeptide that are contacted with a test compound to observe binding, or stimulation or inhibition of a functional response. The functional response of the cells contacted with the test compound is then compared with control cells that were not contacted with the test compound. Such an assay may assess whether the test compound results in a

signal generated by activation of the polypeptide or tetrameric cation channel of the invention, using an appropriate detection system. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist in the presence of the test compound is observed.

A preferred method for identifying an agonist or antagonist compound of a polypeptide or a tetrameric cation channel of the present invention comprises:

(a) contacting a cell expressing on the surface thereof the polypeptide according to the first or second aspect of the invention or the tetrameric cation channel according to the seventh aspect of the invention, the polypeptide or tetrameric cation channel being associated with a second component capable of providing a detectable signal in response to the binding of a compound to the polypeptide or tetrameric cation channel, with a compound to be screened under conditions to permit binding to the polypeptide or tetrameric cation channel; and

(b) determining whether the compound binds to and activates or inhibits the polypeptide or tetrameric cation channel by measuring the level of a signal generated from the interaction of the compound with the polypeptide or tetrameric cation channel.

A further preferred method for identifying an agonist or antagonist of a polypeptide or tetrameric cation channel of the invention comprises:

(a) contacting a cell expressing on the surface thereof the polypeptide or tetrameric cation channel, the polypeptide or tetrameric cation channel being associated with a second component capable of providing a detectable signal in response to the binding of a compound to the polypeptide or tetrameric cation channel, with a compound to be screened under conditions to permit binding to the polypeptide or tetrameric cation channel; and

(b) determining whether the compound binds to and activates or inhibits the polypeptide or tetrameric cation channel by comparing the level of a signal generated from the interaction of the compound with the polypeptide or tetrameric cation channel with the level of a signal in the absence of the compound.

In further preferred embodiments, the general methods that are described above may further comprise conducting the identification of agonist or antagonist in the presence of labelled or unlabelled ligand for the polypeptide or tetrameric cation channel.

In another embodiment of the method for identifying agonist or antagonist compounds of a polypeptide or tetrameric cation channel of the present invention comprises:

determining the inhibition of binding of a ligand to cells which have a polypeptide or tetrameric cation channel of the invention on the surface thereof, or to cell membranes containing such a polypeptide or tetrameric cation channel, in the presence of a candidate compound under conditions to permit binding to the polypeptide or tetrameric cation channel, and determining the amount of ligand bound to the polypeptide or tetrameric cation channel. A compound capable of causing reduction of binding of a ligand is considered to be an agonist or antagonist. Preferably the ligand is labelled.

More particularly, a method of screening for a polypeptide or tetrameric cation channel antagonist or agonist compound comprises the steps of:

(a) incubating a labelled ligand with a whole cell expressing a polypeptide or tetrameric cation channel according to the invention on the cell surface, or a cell membrane containing a polypeptide or tetrameric cation channel of the invention,

(b) measuring the amount of labelled ligand bound to the whole cell or the cell membrane;

(c) adding a candidate compound to a mixture of labelled ligand and the whole cell or the cell membrane of step (a) and allowing the mixture to attain equilibrium;

(d) measuring the amount of labelled ligand bound to the whole cell or the cell membrane after step (c); and

(e) comparing the difference in the labelled ligand bound in step (b) and (d), such that the compound which causes the reduction in binding in step (d) is considered to be an agonist or antagonist.

In certain of the embodiments described above, simple binding assays may be used, in which the adherence of a test compound to a surface bearing the polypeptide or tetrameric cation channel is detected by means of a label directly or indirectly associated with the test compound or in an assay involving competition with a labelled competitor. In another embodiment, competitive drug screening assays may be used, in which neutralising antibodies that are capable of binding the polypeptide or tetrameric cation channel specifically compete with a test compound for binding. In this manner, the antibodies can be used to detect the presence of any test compound that possesses specific binding affinity for the polypeptide or tetrameric cation channel.

Assays may also be designed to detect the effect of added test compounds on the production of mRNA encoding the polypeptide in cells. For example, an ELISA may be constructed that

measures secreted or cell-associated levels of polypeptide using monoclonal or polyclonal antibodies by standard methods known in the art, and this can be used to search for compounds that may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues. The formation of binding complexes between the polypeptide and the compound being tested may then be measured.

Assay methods that are also included within the terms of the present invention are those that involve the use of the genes and polypeptides of the invention in overexpression or ablation assays. Such assays involve the manipulation of levels of these genes/polypeptides in cells and assessment of the impact of this manipulation event on the physiology of the manipulated cells. For example, such experiments reveal details of signalling and metabolic pathways in which the particular genes/polypeptides are implicated, generate information regarding the identities of polypeptides with which the studied polypeptides interact and provide clues as to methods by which related genes and proteins are regulated.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the polypeptide or tetrameric cation channel of interest (see International patent application WO84/03564). In this method, large numbers of different small test compounds are synthesised on a solid substrate, which may then be reacted with the polypeptide or tetrameric cation channel of the invention and washed. One way of immobilising the polypeptide or tetrameric cation channel is to use non-neutralising antibodies. Bound polypeptide or tetrameric cation channel may then be detected using methods that are well known in the art. Purified polypeptide or tetrameric cation channel can also be coated directly onto plates for use in the aforementioned drug screening techniques.

The polypeptide or tetrameric cation channel of the invention may be used to identify membrane-bound or soluble receptors, through standard receptor binding techniques that are known in the art, such as ligand binding and crosslinking assays in which the polypeptide is labelled with a radioactive isotope, is chemically modified, or is fused to a peptide sequence that facilitates its detection or purification, and incubated with a source of the putative receptor (for example, a composition of cells, cell membranes, cell supernatants, tissue extracts, or bodily fluids). The efficacy of binding may be measured using biophysical techniques such as surface plasmon resonance (supplied by Biacore AB, Uppsala, Sweden) and spectroscopy. Binding assays may be used for the purification and cloning of the receptor,

but may also identify agonists and antagonists of the polypeptide, that compete with the binding of the polypeptide or the tetrameric cation channel to its receptor. Standard methods for conducting screening assays are well understood in the art.

The invention also includes a screening kit useful in the methods for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, that are described above.

The invention includes the agonists, antagonists, ligands, receptors, substrates and enzymes, and other compounds which modulate the activity or antigenicity of the polypeptide or tetrameric cation channel of the invention discovered by the methods that are described above.

The invention also provides pharmaceutical compositions comprising a polypeptide, nucleic acid, ligand or compound of the invention in combination with a suitable pharmaceutical carrier. These compositions may be suitable as therapeutic or diagnostic reagents, as vaccines, or as other immunogenic compositions, as outlined in detail below.

According to the terminology used herein, a composition containing a polypeptide, nucleic acid, ligand or compound [X] is "substantially free of" impurities [herein, Y] when at least 85% by weight of the total X+Y in the composition is X. Preferably, X comprises at least about 90% by weight of the total of X+Y in the composition, more preferably at least about 95%, 98% or even 99% by weight.

The pharmaceutical compositions should preferably comprise a therapeutically effective amount of the polypeptide, nucleic acid molecule, ligand, or compound of the invention. The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent needed to treat, ameliorate, or prevent a targeted disease, disorder or condition, to exhibit a detectable therapeutic or preventative effect or to function as a contraceptive. For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, for example, of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

The precise effective amount for a human subject will depend upon the severity of the disease state or of the disorder, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and

tolerance/response to therapy. This amount can be determined by routine experimentation and is within the judgement of the clinician. Generally, an effective dose will be from 0.01 mg/kg to 50 mg/kg, preferably 0.05 mg/kg to 10 mg/kg. Compositions may be administered individually to a patient or may be administered in combination with other agents, drugs or hormones.

A pharmaceutical composition may also contain a pharmaceutically acceptable carrier, for administration of a therapeutic agent. Such carriers include antibodies and other polypeptides, genes and other therapeutic agents such as liposomes, provided that the carrier does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Suitable carriers may be large, slowly metabolised macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers and inactive virus particles.

Pharmaceutically acceptable salts can be used therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulphates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable carriers is available in Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Pharmaceutically acceptable carriers in therapeutic compositions may additionally contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such compositions. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Once formulated, the compositions of the invention can be administered directly to the subject. The subjects to be treated can be animals; in particular, human subjects can be treated.

The pharmaceutical compositions utilised in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal or transcutaneous applications (for example, see WO98/20734), subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, intravaginal or rectal means. Gene guns or hyposprays may also be used to administer the pharmaceutical compositions of the invention. Typically, the therapeutic

compositions may be prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared.

Direct delivery of the compositions will generally be accomplished by injection, subcutaneously, intraperitoneally, intravenously or intramuscularly, or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Dosage treatment may be a single dose schedule or a multiple dose schedule.

If the activity of the polypeptide of the invention is in excess in a particular disease state or disorder, several approaches are available. One approach comprises administering to a subject an inhibitor compound (antagonist) as described above, along with a pharmaceutically acceptable carrier in an amount effective to inhibit the function of the polypeptide, such as by blocking the binding of ligands, substrates, enzymes, receptors, or by inhibiting a second signal, and thereby alleviating the abnormal condition. Preferably, such antagonists are antibodies. Most preferably, such antibodies are chimeric and/or humanised to minimise their immunogenicity, as described previously.

In another approach, soluble forms of the polypeptide that retain binding affinity for the ligand, substrate, enzyme, receptor, in question, may be administered. Typically, the polypeptide may be administered in the form of fragments that retain the relevant portions.

In an alternative approach, expression of the gene encoding the polypeptide can be inhibited using expression blocking techniques, such as the use of antisense nucleic acid molecules (as described above), either internally generated or separately administered. Modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5' or regulatory regions (signal sequence, promoters, enhancers and introns) of the gene encoding the polypeptide. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. *et al.* (1994) In: Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY). The complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes. Such oligonucleotides may be administered or may be generated *in situ* from expression *in vivo*.

In addition, expression of the polypeptide or tetrameric cation channel of the invention may be prevented by using ribozymes specific to the polypeptide's encoding mRNA sequence. Ribozymes are catalytically active RNAs that can be natural or synthetic (see for example Usman, N, *et al.*, Curr. Opin. Struct. Biol (1996) 6(4), 527-33). Synthetic ribozymes can be designed to specifically cleave mRNAs at selected positions thereby preventing translation of the mRNAs into functional polypeptide. Ribozymes may be synthesised with a natural ribose phosphate backbone and natural bases, as normally found in RNA molecules. Alternatively the ribozymes may be synthesised with non-natural backbones, for example, 2'-O-methyl RNA, to provide protection from ribonuclease degradation and may contain modified bases.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of non-traditional bases such as inosine, queosine and butosine, as well as acetyl-, methyl-, thio- and similarly modified forms of adenine, cytidine, guanine, thymine and uridine which are not as easily recognised by endogenous endonucleases.

For treating abnormal conditions related to an under-expression of the polypeptide of the invention and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound that activates the polypeptide, i.e., an agonist as described above, to alleviate the abnormal condition. Alternatively, a therapeutic amount of the polypeptide in combination with a suitable pharmaceutical carrier may be administered to restore the relevant physiological balance of polypeptide.

Gene therapy may be employed to effect the endogenous production of the polypeptide by the relevant cells in the subject. Gene therapy is used to treat permanently the inappropriate production of the polypeptide by replacing a defective gene with a corrected therapeutic gene.

Gene therapy of the present invention can occur *in vivo* or *ex vivo*. *Ex vivo* gene therapy requires the isolation and purification of patient cells, the introduction of a therapeutic gene and introduction of the genetically altered cells back into the patient. In contrast, *in vivo* gene therapy does not require isolation and purification of a patient's cells.

The therapeutic gene is typically "packaged" for administration to a patient. Gene delivery vehicles may be non-viral, such as liposomes, or replication-deficient viruses, such as adenovirus as described by Berkner, K.L., in Curr. Top. Microbiol. Immunol., 158, 39-66 (1992) or adeno-associated virus (AAV) vectors as described by Muzyczka, N., in Curr. Top. Microbiol. Immunol., 158, 97-129 (1992) and U.S. Patent No. 5,252,479. For example, a nucleic acid molecule encoding a polypeptide of the invention may be engineered for expression in a replication-defective retroviral vector. This expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding the polypeptide, such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo* (see Chapter 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in Human Molecular Genetics (1996), T Strachan and A P Read, BIOS Scientific Publishers Ltd).

Another approach is the administration of "naked DNA" in which the therapeutic gene is directly injected into the bloodstream or muscle tissue.

In situations in which the polypeptides or nucleic acid molecules of the invention are disease-causing agents, the invention provides that they can be used in vaccines to raise antibodies against the disease- or disorder-causing agent.

Vaccines according to the invention may either be prophylactic (ie. to prevent infection) or therapeutic (ie. to treat disease after infection). Such vaccines comprise immunising antigen(s), immunogen(s), polypeptide(s), protein(s) or nucleic acid, usually in combination with pharmaceutically-acceptable carriers as described above, which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the antigen or immunogen may be conjugated to a bacterial toxoid, such as a toxoid from diphtheria, tetanus, cholera, H. pylori, and other pathogens.

Since polypeptides may be broken down in the stomach, vaccines comprising polypeptides are preferably administered parenterally (for instance, subcutaneous, intramuscular, intravenous, or intradermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the

recipient, and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents.

The vaccine formulations of the invention may be presented in unit-dose or multi-dose containers. For example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Genetic delivery of antibodies that bind to polypeptides according to the invention may also be effected, for example, as described in International patent application WO98/55607.

The technology referred to as jet injection (see, for example, www.powderject.com) may also be useful in the formulation of vaccine compositions.

A number of suitable methods for vaccination and vaccine delivery systems are described in International patent application WO00/29428.

This invention also relates to the use of nucleic acid molecules according to the present invention as diagnostic reagents. Detection of a mutated form of the gene characterised by the nucleic acid molecules of the invention which is associated with a disease or disorder will provide a diagnostic tool that can add to, or define, a diagnosis of a disease or disorder, or susceptibility to a disease or disorder, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals carrying mutations in the gene may be detected at the DNA level by a variety of techniques.

Nucleic acid molecules for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR, ligase chain reaction (LCR), strand displacement amplification (SDA), or other amplification techniques (see Saiki *et al.*, Nature, 324, 163-166 (1986); Bej, *et al.*, Crit. Rev. Biochem. Molec. Biol., 26, 301-334 (1991); Birkenmeyer *et al.*, J. Virol. Meth., 35, 117-126 (1991); Van Brunt, J., Bio/Technology, 8, 291-294 (1990)) prior to analysis.

In one embodiment, this aspect of the invention provides a method of diagnosing a disease or disorder in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide according to the invention and comparing said level of expression to a control level, wherein a level that is different to said control level is indicative of disease or disorder.

The method may comprise the steps of:

a)contacting a sample of tissue from the patient with a nucleic acid probe under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule of the invention and the probe;

b)contacting a control sample with said probe under the same conditions used in step a);

c)and detecting the presence of hybrid complexes in said samples;

wherein detection of levels of the hybrid complex in the patient sample that differ from levels of the hybrid complex in the control sample is indicative of disease or disorder.

A further aspect of the invention comprises a diagnostic method comprising the steps of:

a)obtaining a tissue sample from a patient being tested for a disease or disorder;

b)isolating a nucleic acid molecule according to the invention from said tissue sample; and

c)diagnosing the patient for a disease or disorder by detecting the presence of a mutation in the nucleic acid molecule which is associated with the disease or disorder.

To aid the detection of nucleic acid molecules in the above-described methods, an amplification step, for example using PCR, may be included.

Deletions and insertions can be detected by a change in the size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridising amplified DNA to labelled RNA of the invention or alternatively, labelled antisense DNA sequences of the invention. Perfectly-matched sequences can be distinguished from mismatched duplexes by RNase digestion or by assessing differences in melting temperatures. The presence or absence of the mutation in the patient may be detected by contacting DNA with a nucleic acid probe that hybridises to the DNA under stringent conditions to form a hybrid double-stranded molecule, the hybrid double-stranded molecule having an unhybridised portion of the nucleic acid probe strand at any portion corresponding to a mutation associated with a disease or disorder; and detecting the presence or absence of an unhybridised portion of the probe strand as an indication of the presence or absence of a disease- or disorder-associated mutation in the corresponding portion of the DNA strand.

Such diagnostics are particularly useful for prenatal and even neonatal testing.

Point mutations and other sequence differences between the reference gene and "mutant" genes can be identified by other well-known techniques, such as direct DNA sequencing or

single-strand conformational polymorphism, (see Orita *et al.*, Genomics, 5, 874-879 (1989)). For example, a sequencing primer may be used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabelled nucleotides or by automatic sequencing procedures with fluorescent-tags. Cloned DNA segments may also be used as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. Further, point mutations and other sequence variations, such as polymorphisms, can be detected as described above, for example, through the use of allele-specific oligonucleotides for PCR amplification of sequences that differ by single nucleotides.

DNA sequence differences may also be detected by alterations in the electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (for example, Myers *et al.*, Science (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton *et al.*, Proc. Natl. Acad. Sci. USA (1985) 85: 4397-4401).

In addition to conventional gel electrophoresis and DNA sequencing, mutations such as microdeletions, aneuploidies, translocations, inversions, can also be detected by *in situ* analysis (see, for example, Keller *et al.*, DNA Probes, 2nd Ed., Stockton Press, New York, N.Y., USA (1993)), that is, DNA or RNA sequences in cells can be analysed for mutations without need for their isolation and/or immobilisation onto a membrane. Fluorescence *in situ* hybridisation (FISH) is presently the most commonly applied method and numerous reviews of FISH have appeared (see, for example, Trachuck *et al.*, Science, 250, 559-562 (1990), and Trask *et al.*, Trends, Genet., 7, 149-154 (1991)).

In another embodiment of the invention, an array of oligonucleotide probes comprising a nucleic acid molecule according to the invention can be constructed to conduct efficient screening of genetic variants, mutations and polymorphisms. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M.Chee *et al.*, Science (1996), Vol 274, pp 610-613).

In one embodiment, the array is prepared and used according to the methods described in PCT application WO95/11995 (Chee *et al*); Lockhart, D. J. *et al.* (1996) Nat. Biotech. 14: 1675-1680); and Schena, M. *et al.* (1996) Proc. Natl. Acad. Sci. 93: 10614-10619). Oligonucleotide

pairs may range from two to over one million. The oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass slide or any other suitable solid support. In another aspect, an oligonucleotide may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application W095/25116 (Baldeschweiler *et al*). In another aspect, a "gridded" array analogous to a dot (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. An array, such as those described above, may be produced by hand or by using available devices (slot blot or dot blot apparatus), materials (any suitable solid support), and machines (including robotic instruments), and may contain 8, 24, 96, 384, 1536 or 6144 oligonucleotides, or any other number between two and over one million which lends itself to the efficient use of commercially-available instrumentation.

In addition to the methods discussed above, diseases and disorders may be diagnosed by methods comprising determining, from a sample derived from a subject, an abnormally decreased or increased level of polypeptide or mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridisation methods.

Assay techniques that can be used to determine levels of a polypeptide of the present invention in a sample derived from a host are well-known to those of skill in the art and are discussed in some detail above (including radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays). This aspect of the invention provides a diagnostic method which comprises the steps of: (a) contacting a ligand as described above with a biological sample under conditions suitable for the formation of a ligand-polypeptide complex; and (b) detecting said complex.

Protocols such as ELISA, RIA, and FACS for measuring polypeptide levels may additionally provide a basis for diagnosing altered or abnormal levels of polypeptide expression. Normal or standard values for polypeptide expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably humans, with antibody to the polypeptide under conditions suitable for complex formation The amount of standard complex formation may be quantified by various methods, such as by photometric means.

Antibodies which specifically bind to a polypeptide of the invention may be used for the diagnosis of disorders or diseases characterised by expression of the polypeptide, or in assays to monitor patients being treated with the polypeptides, nucleic acid molecules, ligands and other compounds of the invention or who are using the same as a contraceptive agent. Antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for the polypeptide include methods that utilise the antibody and a label to detect the polypeptide in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labelled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules known in the art may be used, several of which are described above.

Quantities of polypeptide expressed in subject, control and disease/disorder samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing a disease or disorder. Diagnostic assays may be used to distinguish between absence, presence, and excess expression of polypeptide and to monitor regulation of polypeptide levels during therapeutic intervention. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

A diagnostic kit of the present invention may comprise:

(a) a nucleic acid molecule of the present invention;

(b) a polypeptide of the present invention;

(c) a tetrameric cation channel of the present invention; or

(d) a ligand of the present invention.

In one aspect of the invention, a diagnostic kit may comprise a first container containing a nucleic acid probe that hybridises under stringent conditions with a nucleic acid molecule according to the invention; a second container containing primers useful for amplifying the nucleic acid molecule; and instructions for using the probe and primers for facilitating the diagnosis of a disease or disorder. The kit may further comprise a third container holding an agent for digesting unhybridised RNA.

In an alternative aspect of the invention, a diagnostic kit may comprise an array of nucleic acid molecules, at least one of which may be a nucleic acid molecule according to the invention.

To detect a polypeptide according to the invention, a diagnostic kit may comprise one or more antibodies that bind to a polypeptide according to the invention; and a reagent useful for the detection of a binding reaction between the antibody and the polypeptide.

To detect a tetrameric cation channel according to the invention, a diagnostic kit may comprise one or more antibodies that bind to a tetrameric cation channel according to the invention; and a reagent useful for the detection of a binding reaction between the antibody and the tetrameric cation channel.

Such kits will be of use in diagnosing a disease or disorder or susceptibility to a disease or disorder, including, but not limited to, disorders such as infertility, neurological disorders or cardiovascular disorders, autoimmune disease or stroke, as well as other conditions in which cation channels are implicated.

Various aspects and embodiments of the present invention will now be described in more detail by way of example, with particular reference to the INPIONCH05 and INPIONCH06 polypeptides.

It will be appreciated that modification of detail may be made without departing from the scope of the invention.

Brief description of the Figures

Figure 1: Results from BLAST against NCBI non-redundant database using the full length INPIONCH05 polypeptide sequence;

Figure 2: Alignment generated by BLAST between the full-length INPIONCH05 polypeptide sequence and the sequence of the closest known protein, (L01983) sodium channel alpha subunit from *Homo sapiens*;

Figure 3: Results from BLAST against NCBI non-redundant database using the masked full length INPIONCH05 sequence. By 'masked' it is meant that where there exists a long row of hydrophobic residues, these have been replaced ('masked') by a row of Xs;

Figure 4: Alignment generated by BLAST between the masked INPIONCH05 polypeptide sequence and the sequence of the closest known protein, (XM_083858.1) cation channel of sperm from *Homo sapiens*;

Figure 5: Conserved domain search results: PFAM sequence motifs producing significant alignments to the amino acid sequence of INPIONCH05;

Figure 6: InterproScan results for INPIONCH05;

Figure 7: Ion Trans Domain Multiple Alignment between CatSper family members showing transmembrane regions highlighted in red, and TxDxW ion_selectivity region boxed;

Figure 8: INPIONCH05 topology prediction using the TMHMM transmembrane domain prediction programme;

Figure 9: Coiled-coil prediction from Lupas Coiled coils prediction program (Lupas A. *et al.*, Predicting Coiled Coils from Protein Sequences. Science, 252, 1162-1164 (1991));

Figure 10: Sequences of CatSper1;

Figure 11: Sequences of CatSper2;

Figure 12: Sequences of CatSper3;

Figure 13: Sequences of CatSper2 human splice variants;

Figure 14: ENSEMBL predicted amino acid transcript: ENST00000294438; and

Figure 15: Normalised expression of INPIONCH05 in 22 samples derived from 18 normal human tissues.

Examples

Example 1: INPIONCH05 and INPIONCH06

The polypeptide sequence derived from combining SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18 and SEQ ID NO: 20 and SEQ ID NO:22, which represents the translation of consecutive exons from INPIONCH05, was used as a BLAST query against the NCBI non-redundant Sequence database. The top twelve matches are shown in Figure 1. The majority of the top matches are to putative or known cation channel subunits thus reinforcing the functional prediction of the polypeptides of the invention.

Figure 2 shows the alignment of the INPIONCH05 query sequence to the sequence of the highest matching known protein, sodium channel alpha subunit from *Homo sapiens* (L01983).

The eleventh highest match is to a sperm ion channel from *Mus. musculus* (AF407332).

The polypeptide sequence derived from combining SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18 and SEQ ID NO: 20 and SEQ ID NO:22, which represents the translation of consecutive exons from INPIONCH05 and in which low-complexity or coiled-coil regions of the query sequence have been 'masked' out with X's was used as a BLAST query against the NCBI non-redundant Sequence database. Low-complexity or coiled-coil regions can match many sequences in the database, thus causing "matrix drift", permitting large numbers of un-related sequences to be dragged into the sequence profile of your query. It is for this reason that these region are masked. The top ten matches are shown in Figure 3. The majority of the top matches are to putative or known cation channel subunits.

When the masked sequence is used in the BLAST, as opposed to the unmasked sequence (results shown in Figure 1) the top match to a known protein is now the second highest match. This alignment is to a cation channel of sperm from *Homo sapiens* (XM_083858) and is shown in Figure 4.

The amino acid sequence of INPIONCH05 (SEQ ID NO:22) was submitted into the 3D-PSSM (position specific scoring matrix) search tool database (http://www.sbg.bio.ic.ac.uk/~3dpssm/) in an attempt to predict its 3-dimensional structure and possible function. The most significant alignment was to the amino acid sequence of pfam00520 ion_trans, Ion transport protein and is shown in Figure 5. This family contains Na^+, K^+ and Ca^{2+} channels. The proteins in this family have 6 transmembrane helices in which the last two helices flank a loop which determines ion selectivity. In some sub-families, e.g. Na^+ channels, the domain is repeated four times, whereas in others, e.g. K^+ channels, the protein forms as a tetramer in the membrane.

Figure 6 shows the InterproScan results for INPIONCH05 (SEQ ID NO:22). InterProScan is a tool that combines different protein signature recognition methods native to the InterPro member databases into one resource with look up of corresponding InterPro and GO annotation. The results show that SEQ ID NO:22 is a cation channel. They also give a clue as to the cation specificity of SEQ ID NO:22; it is not a K^+ channel subunit but may be either a Ca^{2+} or a Na^+ channel subunit.

The high level of sequence identity between the ion_trans domains of INPIONCH05, INPIONCH06 and other members of the CatSper family such as human and mouse CatSper1, human and mouse CatSper2, human CatSper2 splice variants and human and mouse CatSper3 is shown in Figure 7. The boxed region shows the highly conserved TxDxW ion_selectivity region which is conserved between all the sequences. The highly conserved aspartate (D) in this region is critical for calcium ion selectivity (Quill,T.A. *et al.*, A voltage-gated ion channel expressed specifically in spermatozoa. PNAS, Vol. 98(22) 12527-12531 (2001)). This therefore provides further evidence that the INPIONCH05 and INPIONCH06 polypeptides of the present invention are calcium-specific cation channels or channel subunits.

The topology prediction of INPIONCH05 (SEQ ID NO:22) using the TMHMM programme for prediction of transmembrane helices in proteins shows that the protein adopts a six transmembrane helical structure. This is therefore similar to the six-transmembrane spanning domains of the alpha subunits of the Na^+ and Ca^{2+} voltage-gated channels.

The presence of coiled coil domains towards the C-terminal of each of CatSper1, CatSper2, INPIONCH05 and CatSper3 is shown in Figure 9. These domains may serve as sites of protein interactions between identical or different channel subunits to form homo- or hetero-tetrameric channels. It is possible that these coiled coil domains may also form a site for interaction of the subunits with other accessory proteins.

Figure 14 gives the sequence of an ENSEMBL predicted amino acid transcript which bears some level of sequence identity to the polypeptides of the present invention. However, unlike the sequence of SEQ ID NO:22, this ENSEMBL predicted sequence only has around 180 amino acids and also has a different position of one of its exons.

Example 2: Data for INPIONCH05

Materials and methods

Taqman-RT-PCR: The TaqMan 3'-5' exonuclease assay signals the formation of PCR amplicons by a process involving the nucleolytic degradation of a double-labelled fluorogenic probe that hybridises to the target template at a site between the two primer recognition sequences (see US Patent 5,876,930). The ABI Prism 7700 automates the detection and quantitative measurement of these signals, which are stoichiometrically related to the quantities of amplicons produced, during each cycle of amplification. In addition to providing

substantial reductions in the time and labour requirements for PCR analyses, this technology permits simplified and potentially highly accurate quantification of target sequences in the reactions.

Taqman RT-PCR was carried out using 15ng of INPIONCH05 cDNA using primers/probes specific for INPIONCH05 cDNA and 18s rRNA. Within each experiment, a standard curve for target and internal control for a typical tissue sample was calculated, using between 25ng to 0.39ng of cDNA template. Cycle threshold (Ct) determinations (i.e. non-integer calculations of the number of cycles required for reporter dye fluorescence resulting from the synthesis of PCR products to become significantly higher than background fluorescence levels) were performed by the instrument for each reaction using default parameters. Using linear regression analysis of the standard curves, the Ct values were used to calculate the amount of actual starting target or 18s cDNA in each test sample.

The levels of target cDNA in each sample were normalised to the level of expression of target cDNA in a comparative sample, in this case, heart. The levels of 18s cDNA (the internal control) in each sample were also normalised to the level of expression of 18s cDNA in heart. The expression levels of INPIONCH05 cDNA were then normalised to the expression levels of 18s cDNA.

RNA samples: Human RNA prepared from non-diseased organs was purchased from either Ambion Europe (Huntingdon, UK) or Clontech (BD, Franklin Lakes, NJ).

Oligonucleotide design: Oligonucleotide primers and probes were designed using Primer Express software (Applied Biosystems, Foster City CA) with a GC-content of 40-60%, no G-nucleotide at the 5'-end of the probe, and no more than 4 contiguous Gs. Each primer and probe was analysed using BLAST® (Basic Local Alignment Search Tool, Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.: J Mol Biol 1990 Oct 5;215(3):403-10). The results confirm that each oligonucleotide recognised the target sequence with a specificity of >3 bp when compared to other known cDNAs or genomic sequence represented in the Unigene and GoldenPath publicly available databases.

The sequence of the primers and probes were: INPIONCH05Fwd:

AAGGACATCCGCCAGATGTC;

INPIONCH05Probe: CAACAGCAAGACTTGCTCAGTGCGCT and

INPIONCH05Rev: GGCACACCTTTTCCATGCTAA.

18s pre-optimised primers and probe were purchased from Applied Biosystems, Foster City, CA.

In the fluorescence studies, the probes were covalently conjugated with a fluorescent reporter dye (e.g. 6carboxy-fluorescein [FAM]; Xem = 518nm) and a fluorescent quencher dye (6carboxytetram-ethyl-rhodamine [TAMRA]; Mem = 582nm) at the most 5' and most 3' base, respectively. Primers were obtained from Sigma Genosys, UK and probes are obtained from Eurogentec, Belgium. Primer/probe concentrations were titrated in the range of 50nM to 900nM and optimal concentrations for efficient PCR reactions were determined. Optimal primer and probe concentrations varied between 100nM and 900nM depending on the target gene that was to be amplified.

cDNA reaction: cDNA was prepared using components from Applied Biosystems, Foster City CA. 50µl reactions were prepared in 0.5ml RNase-free tubes. Reactions contained 500ng total RNA, 1x reverse transcriptase buffer; 5.5mM $MgCl_2$, 1mM dNTPs, 2.5µl random hexamers; 20U RNase inhibitor and 62.5U reverse transcriptase.

PCR reactions: 25µl reactions were prepared in 0.5 ml thin-walled optical grade PCR 96 well plates (Applied Biosystems, Foster City CA). The reactions contained: 1x final concentration of TaqMan Universal Master Mix (a proprietary mixture of AmpliTaq Gold DNA polymerase, AmpEraseX UNG, dNTPs with UTP, passive reference dye and optimised buffer components, Applied Biosystems, Foster City CA), 100nM Taqman probe, 900nM forward primer, 900nM reverse primer and 15ng of cDNA template.

Performance of Assay: Standard procedures for the operation of the ABI Prism 7700 or similar detection system were used. For example, for the ABI Prism 7700, these included the use of all default program settings with the exception of the reaction volume, which was changed from 50 to 25 ul. Thermal cycling conditions consisted of two min at 50°C, 10 min at 95°C, followed by 40 cycles of 15 sec at 95°C and 1 min at 60°C. Cycle threshold (Ct) determinations (i.e. non-integer calculations of the number of cycles required for reporter dye fluorescence resulting from the synthesis of PCR products to become significantly higher than background fluorescence levels) were automatically performed by the instrument for each reaction using default parameters. Assays for target sequences and ribosomal 18s (reference) sequences in the same cDNA samples were performed in separate reaction tubes.

Results

Taqman RT-PCR quantitation was used to determine the tissue expression of the INPIONCH05 ion channel. The primer/probe set recognised a sequence within exon 9 of the INPIONCH05 coding sequence.

Figure 15 shows normalised expression of INPIONCH05 in 22 samples derived from 18 normal human tissues and represents the fold expression of normalised target sequence relative to the level of expression of the target sequence in the comparative sample (heart), which is set arbitrarily to 1. Each sample was quantified in 3 individual experiments. Figure 15 also shows the mean ± SEM for the multiple experiments.

The Taqman RT-PCR reaction was also carried out on cDNA made in the absence of reverse transcriptase enzyme. No signal was seen in these reactions (data not shown), indicating that results shown in Figure 15 are not affected by genomic contamination of the RNA sample and that the region of INPIONCH05 detected is present in the cDNA.

The results indicate that INPIONCH05 shows a much higher level of expression in testis than in lung and placenta. The level of expression in the other tissues tested was less than in these three tissues. These results provide evidence for the finding that INPIONCH05 is a testis-specific cation channel of the CatSper family.

Sequences

Note: for amino acids encoded by exon-exon junctions, the amino acid will be assigned to the more 5' exon.

SEQ ID: 1 INPIONCH05 exon 1 nucleotide sequence

```
  1  ATGAGGGATA ATGAAAAGGC CTGGTGGCAG CAATGGACCT CCCATACAGG CCTCGAGGGG
 61  TGGGGCGGGA CTCAGGAGGA CCGTATGGGG TTTGGAGGGG CAGTAGCTGC ACTGAGGGGC
121  CGCCCCTCTC CCCTGCAGAG TACCATTCAC GAGTCCTACG GTCGGCCAGA GGAGCAAGTG
181  CTCATCAACC GCCAGGAAAT CACGAACAAA GCG
```

SEQ ID: 2 INPIONCH05 exon 1 amino acid sequence

```
  1  MRDNEKAWWQ QWTSHTGLEG WGGTQEDRMG FGGAVAALRG RPSPLQSTIH ESYGRPEEQV
 61  LINRQEITNK A
```

SEQ ID: 3 INPIONCH05 exon 2 nucleotide sequence

```
  1  GACGCCTGGG ACATGCAGGA GTTCATCACT CACATGTACA TCAAGCAGCT GCTCCGACAC
 61  CCCGCCTTCC AACTGCTGCT GGCCCTGCTG CTGGTGATCA ATGCCATCAC CATCGCTCTC
121  CGTACCAACT CCTACCTGGA CCAG
```

SEQ ID: 4 INPIONCH05 exon 2 amino acid sequence

```
  1  DAWDMQEFIT HMYIKQLLRH PAFQLLLALL LVINAITIAL RTNSYLDQ
```

SEQ ID: 5 INPIONCH05 exon 3 nucleotide sequence

```
  1  AAACACTATG AGTTGTTCTC TACCATAGAT GACATTGTGC TGACCATCCT TCTTTGTGAG
 61  GTTCTCCTTG GCTGGCTCAA TGGCTTCTGG ATTTTCTGGA AG
```

SEQ ID: 6 INPIONCH05 exon 3 amino acid sequence

```
  1  KHYELFSTID DIVLTILLCE VLLGWLNGFW IFWK
```

SEQ ID: 7 INPIONCH05 exon 4 nucleotide sequence

```
  1  GACGGCTGGA ACATCCTCAA CTTCATTATC GTCTTTATCT TGCTCTTGCG GTTCTTCATT
 61  AATGAAATCA ATATTCCCTC CATCAACTAC ACTCTCAG
```

SEQ ID: 8 INPIONCH05 exon 4 amino acid sequence

```
  1  DGWNILNFII VFILLLRFFI NEINIPSINY TLR
```

SEQ ID: 9 INPIONCH05 exon 5 nucleotide sequence

```
  1  GGCGCTTCGT CTGGTGCATG TGTGCATGGC GGTGGAGCCC CTCGCCCGGA TCATCCGCGT
 61  CATCCTGCAG TCGGTGCCTG ACATGGCCAA TATCATGGTC CTCATCCTCT TCTTCATGCT
121  G
```

SEQ ID: 10 INPIONCH05 exon 5 amino acid sequence

```
  1  ALRLVHVCMA VEPLARIIRV ILQSVPDMAN IMVLILFFML
  2
```

SEQ ID: 11 INPIONCH05 exon 6 nucleotide sequence

```
  1  GTTTTTTCCG TGTTTGGAGT AACACTCTTT GGTGCATTCG TGCCCAAGCA TTTCCAGAAC
 61  ATACAGGTTG CGCTGTACAC CCTCTTCATC TGCATCACCC AGGACGGCTG GGTGGACATC
121  TACAGTGACT TCCA
```

SEQ ID: 12 INPIONCH05 exon 6 amino acid sequence

```
  1  VFSVFGVTLF GAFVPKHFQN IQVALYTLFI CITQDGWVDI YSDFQ
```

SEQ ID: 13 INPIONCH05 exon 7 nucleotide sequence

```
  1  GACAGAGAAG AGGGAATATG CAATGGAGAT TGGGGGTGCC ATCTACTTTA CCATCTTCAT
 61  CACCATCGGT GCCTTCATTG GCATCAACCT GTTCGTCATC GTGGTGACCA CCAACCTGGA
121  GCAAATGATG AAGGCAGGAG AGCAGGGACA ACAGCAACGA ATAACCTTTA GTGAG
```

SEQ ID: 14 INPIONCH05 exon 7 amino acid sequence

```
  1  TEKREYAMEI GGAIYFTIFI TIGAFIGINL FVIVVTTNLE QMMKAGEQGQ QQRITFSE
```

SEQ ID: 15 INPIONCH05 exon 8 nucleotide sequence

```
  1  ACAGGCGCAG AGGAAGAGGA GGAGAATGAC CAGCTGCCAC TGGTGCATTG TGTGGTCGCC
 61  CGCTCGGAGA AATCTGGTCT CCTCCAGGAA CCCCTTGCGG GAGGCCCCCT GTCGAACCTC
121  TCAGAAAACA CGTGTGACAA CTTTTGCTTG GTGCTTGAGG CAATACAGGA GAACCTGAGG
181  CAGTACAAGG AGATCCGAGA TGAACTCAAC AT
```

SEQ ID: 16 INPIONCH05 exon 8 amino acid sequence

```
  1  TGAEEEEEND QLPLVHCVVA RSEKSGLLQE PLAGGPLSNL SENTCDNFCL VLEAIQENLR
 61  QYKEIRDELN M
```

SEQ ID: 17 INPIONCH05 exon 9 nucleotide sequence

```
  1  GATTGTGGAG GAGGTGCGCG CAATCCGCTT CAACCAGGAG CAGGAGTCAG AGGTGTTGAA
 61  CAGGCGCTCG TCGACGAGCG GGTCGTTGGA GACTACGTCA TCCAAGGACA TCCGCCAGAT
121  GTCTCAACAG CAAGACTTGC TCAGTGCGCT CGTTAGCATG GAAAAGGTGT GCCTTCCTTC
181  TCCTACCCAA TGG
```

SEQ ID: 18 INPIONCH05 exon 9 amino acid sequence

```
   1  IVEEVRAIRF NQEQESEVLN RRSSTSGSLE TTSSKDIRQM SQQQDLLSAL VSMEKVCLPS
  61  PTQW
```

SEQ ID: 19 INPIONCH05 exon 10 nucleotide sequence

```
   1  ACCTATGAAG CACTTCTCTG GGAGGCCTGG CTGGGGAAGC TGGTGAGACT GGAGGAGGTG
  61  AGTAACGGAG GCTGCCATAT CGGGGGCAGT GGGGGATTAA CAAGAAGACC TCTTGGTTTG
 121  CAGGTTCATG ACTCTAGCTC ACAAATACTC CTTAAAAAAC ACAAGAGCAG CCACTGA
```

SEQ ID: 20 INPIONCH05 exon 10 amino acid sequence

```
   1  TYEALLWEAW LGKLVRLEEV SNGGCHIGGS GGLTRRPLGL QVHDSSSQIL LKKHKSSH
```

SEQ ID: 21 INPIONCH05 nucleotide sequence

```
   1  ATGAGGGATA ATGAAAAGGC CTGGTGGCAG CAATGGACCT CCCATACAGG CCTCGAGGGG
  61  TGGGGCGGGA CTCAGGAGGA CCGTATGGGG TTTGGAGGGG CAGTAGCTGC ACTGAGGGGC
 121  CGCCCCTCTC CCCTGCAGAG TACCATTCAC GAGTCCTACG GTCGGCCAGA GGAGCAAGTG
 181  CTCATCAACC GCCAGGAAAT CACGAACAAA GCGGACGCCT GGGACATGCA GGAGTTCATC
 241  ACTCACATGT ACATCAAGCA GCTGCTCCGA CACCCCGCCT TCCAACTGCT GCTGGCCCTG
 301  CTGCTGGTGA TCAATGCCAT CACCATCGCT CTCCGTACCA ACTCCTACCT GGACCAGAAA
 361  CACTATGAGT TGTTCTCTAC CATAGATGAC ATTGTGCTGA CCATCCTTCT TTGTGAGGTT
 421  CTCCTTGGCT GGCTCAATGG CTTCTGGATT TTCTGGAAGG ACGGCTGGAA CATCCTCAAC
 481  TTCATTATCG TCTTTATCTT GCTCTTGCGG TTCTTCATTA ATGAAATCAA TATTCCCTCC
 541  ATCAACTACA CTCTCAGGGC GCTTCGTCTG GTGCATGTGT GCATGGCGGT GGAGCCCCTC
 601  GCCCGGATCA TCCGCGTCAT CCTGCAGTCG GTGCCTGACA TGGCCAATAT CATGGTCCTC
 661  ATCCTCTTCT TCATGCTGGT TTTTTCCGTG TTTGGAGTAA CACTCTTTGG TGCATTCGTG
 721  CCCAAGCATT TCCAGAACAT ACAGGTTGCG CTGTACACCC TCTTCATCTG CATCACCCAG
 781  GACGGCTGGG TGGACATCTA CAGTGACTTC CAGACAGAGA AGAGGGAATA TGCAATGGAG
 841  ATTGGGGGTG CCATCTACTT TACCATCTTC ATCACCATCG GTGCCTTCAT TGGCATCAAC
 901  CTGTTCGTCA TCGTGGTGAC CACCAACCTG GAGCAAATGA TGAAGGCAGG AGAGCAGGGA
 961  CAACAGCAAC GAATAACCTT TAGTGAGACA GGCGCAGAGG AAGAGGAGGA GAATGACCAG
1021  CTGCCACTGG TGCATTGTGT GGTCGCCCGC TCGGAGAAAT CTGGTCTCCT CCAGGAACCC
1081  CTTGCGGGAG GCCCCCTGTC GAACCTCTCA GAAAACACGT GTGACAACTT TTGCTTGGTG
1141  CTTGAGGCAA TACAGGAGAA CCTGAGGCAG TACAAGGAGA TCCGAGATGA ACTCAACATG
1201  ATTGTGGAGG AGGTGCGCGC AATCCGCTTC AACCAGGAGC AGGAGTCAGA GGTGTTGAAC
1261  AGGCGCTCGT CGACGAGCGG GTCGTTGGAG ACTACGTCAT CCAAGGACAT CCGCCAGATG
1321  TCTCAACAGC AAGACTTGCT CAGTGCGCTC GTTAGCATGG AAAAGGTGTG CCTTCCTTCT
1381  CCTACCCAAT GGACCTATGA AGCACTTCTC TGGGAGGCCT GGCTGGGGAA GCTGGTGAGA
1441  CTGGAGGAGG TGAGTAACGG AGGCTGCCAT ATCGGGGGGA GTGGGGGATT AACAAGAAGA
1501  CCTCTTGGTT TGCAGGTTCA TGACTCTAGC TCACAAATAC TCCTTAAAAA ACACAAGAGC
1561  AGCCACTGA
```

SEQ ID: 22 INPIONCH05 amino acid sequence

```
   1  MRDNEKAWWQ QWTSHTGLEG WGGTQEDRMG FGGAVAALRG RPSPLQSTIH ESYGRPEEQV
  61  LINRQEITNK ADAWDMQEFI THMYIKQLLR HPAFQLLLAL LLVINAITIA LRTNSYLDQK
 121  HYELFSTIDD IVLTILLCEV LLGWLNGFWI FWKDGWNILN FIIVFILLLR FFINEINIPS
 181  INYTLRALRL VHVCMAVEPL ARIIRVILQS VPDMANIMVL ILFFMLVFSV FGVTLFGAFV
 241  PKHFQNIQVA LYTLFICITQ DGWVDIYSDF QTEKREYAME IGGAIYFTIF ITIGAFIGIN
 301  LFVIVVTTNL EQMMKAGEQG QQQRITFSET GAEEEEENDQ LPLVHCVVAR SEKSGLLQEP
 361  LAGGPLSNLS ENTCDNFCLV LEAIQENLRQ YKEIRDELNM IVEEVRAIRF NQEQESEVLN
 421  RRSSTSGSLE TTSSKDIRQM SQQQDLLSAL VSMEKVCLPS PTQWTYEALL WEAWLGKLVR
 481  LEEVSNGGCH IGGSGGLTRR PLGLQVHDSS SQILLKKHKS SH
```

SEQ ID: 23 INPIONCH06 exon 1 nucleotide sequence

```
  1  ATGTCTGAAA AACACAAGTG GTGGCAGCAG GTGGAGAACA TCGACATCAC ACACCTGGGC
 61  CCTAAGGTAG AGAGCCTCGT CGTGGCGCTG CCGCCGAGCT GGCATCCTAG G
```

SEQ ID: 24 INPIONCH06 exon 1 amino acid sequence

```
  1  MSEKHKWWQQ VENIDITHLG PKVESLVVAL PPSWHPR
```

SEQ ID: 25 INPIONCH06 exon 2 nucleotide sequence

```
  1  AGAAAAGCCT ATGAACTCCT GGGTCGGCAT GAGGAGCAAG TGCTCATCAA CCGCAGAGAT
 61  GTCATGGAGA AGAAG
```

SEQ ID: 26 INPIONCH06 exon 2 amino acid sequence

```
  1  RKAYELLGRH EEQVLINRRD VMEKK
```

SEQ ID: 27 INPIONCH06 exon 3 nucleotide sequence

```
  1  GATGCCTGGG ATGTACAGGA ATTCATCACT CAAATGTATA TCAAGCAGTT GCTCCGCCAT
 61  CCGGCCTTCC AGCTGCTGCT GGCCTTTCTG CTGCTGTCCA ACGCTATCAC CATTGCCCTT
121  CGCACCAACT CTTATCTCGG TCAG
```

SEQ ID: 28 INPIONCH06 exon 3 amino acid sequence

```
  1  DAWDVQEFIT QMYIKQLLRH PAFQLLLAFL LLSNAITIAL RTNSYLGQ
```

SEQ ID: 29 INPIONCH06 exon 4 nucleotide sequence

```
  1  AAACATTACG AGCTATTCTC GACCATAGAT GACATTGTGT TGACGATCCT TATCTGCGAG
 61  GTTCTGCTTG GTTGGCTTAA CGGCTTCTGG ATTTTCTGGA AG
```

SEQ ID: 30 INPIONCH06 exon 4 acid sequence

```
  1  KHYELFSTID DIVLTILICE VLLGWLNGFW IFWK
```

SEQ ID: 31 INPIONCH06 exon 5 nucleotide sequence

```
  1  GATGGCTGGA ATATCCTCAA CTTCGCAATT GTCTTTATCT TGTTTATGGG GTTCTTCATA
 61  AAACAACTTG ACATGGTTGC CATCACCTAC CCTCTCAG
```

SEQ ID: 32 INPIONCH06 exon 5 amino acid sequence

```
  1  DGWNILNFAI VFILFMGFFI KQLDMVAITY PLR
```

SEQ ID: 33 INPIONCH06 exon 6 nucleotide sequence

```
  1  GGTGCTCCGG CTGGTGCATG TGTGTATGGC GGTGGAACCC CTGGCCAGAA TCATCAAGGT
 61  TATCCTGCAG TCGATGCCAG ACTTGGCTAA TGTCATGGCT CTCATCCTCT TCTTCATGCT
121  G
```

SEQ ID: 34 INPIONCH06 exon 6 amino acid sequence

```
  1  VLRLVHVCMA VEPLARIIKV ILQSMPDLAN VMALILFFML
```

SEQ ID: 35 INPIONCH06 exon 7 nucleotide sequence

```
  1  GTATTCTCTG TGTTTGGGGT CACGCTCTTC GGTGCATTTG TGCCCAAGCA TTTCCAGAAC
 61  ATGGGGGTTG CCCTGTACAC GCTCTTCATC TGCATCACTC AGGATGGATG GCTGGACATC
121  TACACTGACT TCCA
```

SEQ ID: 36 INPIONCH06 exon 7 amino acid sequence

```
  1  VFSVFGVTLF GAFVPKHFQN MGVALYTLFI CITQDGWLDI YTDFQ
```

SEQ ID: 37 INPIONCH06 exon 8 nucleotide sequence

```
  1  GATGGATGAA AGAGAGTACG CGATGGAGGT CGGGGGCGCC ATCTACTTTG CCGTCTTTAT
 61  CACCCTCGGT GCCTTCATTG GTCTCAACTT GTTCGTCGTC GTGGTGACCA CAAACCTGGA
121  ACAAATGATG AAGACCGGCG AGGAAGAGGG ACACCTGAAC ATAAAGTTTA CTGAG
```

SEQ ID: 38 INPIONCH06 exon 8 amino acid sequence

```
  1  MDEREYAMEV GGAIYFAVFI TLGAFIGLNL FVVVVTTNLE QMMKTGEEEG HLNIKFTE
```

SEQ ID: 39 INPIONCH06 exon 9 nucleotide sequence

```
  1  ACAGAAGAGG ATGAGGACTG GACCGACGAG CTGCCACTGG TGCATTGTAC AGAGGCCCGC
 61  AAGGATACTT CCACTGTCCC CAAGGAACCA CTGGTTGGGG GCCCCCTGAG TAACCTCACA
121  GAAAAGACCT GCGATAACTT CTGCTTGGTG CTTGAAGCAA TACAGGAGAA CTTGATGGAG
181  TACAAAGAGA TCCGAGAGGA ACTCAACAT
```

SEQ ID: 40 INPIONCH06 exon 9 amino acid sequence

```
  1  TEEDEDWTDE LPLVHCTEAR KDTSTVPKEP LVGGPLSNLT EKTCDNFCLV LEAIQENLME
 61  YKEIREELNM
```

SEQ ID: 41 INPIONCH06 exon 10 nucleotide sequence

```
  1  GATCGTGGAG GAAGTGTCCT CCATCCGGTT CAACCAGGAG CAGCAAAATG TGATCCTACA
 61  CAAGTATACC TCCAAAAGCG CCACCTTCCT AAGCGAGCCC CCAGAAGGGG CTAACAAGCA
121  AGACTTGATC ACTGCGCTGG TCAGCAGGGA AAAG
```

SEQ ID: 42 INPIONCH06 exon 10 amino acid sequence

```
1  IVEEVSSIRF NQEQQNVILH KYTSKSATFL SEPPEGANKQ DLITALVSRE K
```

SEQ ID: 43 INPIONCH06 exon 11 nucleotide sequence

```
  1  AGTAGAACGG GTCCTGGGGA TGGCTACACT CTCCAGTTTA GATTGTTACT AAGCATGCTA
 61  GAGTTACTCG GGGTGGAGCA GCCGTGGAGG CAGGAGGAGG AAGAAGGAGC AGGCTGTCCC
121  ATTAAGGTAA GCAGAGTGAC AAAAAGCTCT CTTGGTTTGC AGGTGTCTGA TTCTAACATA
181  AACATGGTTA ACAAACACAA GTTCAGCCAC TGA
```

SEQ ID: 44 INPIONCH06 exon 11 amino acid sequence

```
1  SRTGPGDGYT LQFRLLLLSML ELLGVEQPWR QEEEEGAGCP IKVSRVTKSS LGLQVSDSNI
61  NMVNKHKFSH
```

SEQ ID: 45 INPIONCH06 nucleotide sequence

```
   1  ATGTCTGAAA AACACAAGTG GTGGCAGCAG GTGGAGAACA TCGACATCAC ACACCTGGGC
  61  CCTAAGGTAG AGAGCCTCGT CGTGGCGCTG CCGCCGAGCT GGCATCCTAG GAGAAAAGCC
 121  TATGAACTCC TGGGTCGGCA TGAGGAGCAA GTGCTCATCA ACCGCAGAGA TGTCATGGAG
 181  AAGAAGGATG CCTGGGATGT ACAGGAATTC ATCACTCAAA TGTATATCAA GCAGTTGCTC
 241  CGCCATCCGG CCTTCCAGCT GCTGCTGGCC TTTCTGCTGC TGTCCAACGC TATCACCATT
 301  GCCCTTCGCA CCAACTCTTA TCTCGGTCAG AAACATTACG AGCTATTCTC GACCATAGAT
 361  GACATTGTGT TGACGATCCT TATCTGCGAG GTTCTGCTTG GTTGGCTTAA CGGCTTCTGG
 421  ATTTTCTGGA AGGATGGCTG GAATATCCTC AACTTCGCAA TTGTCTTTAT CTTGTTTATG
 481  GGGTTCTTCA TAAAACAACT TGACATGGTT GCCATCACCT ACCCTCTCAG GGTGCTCCGG
 541  CTGGTGCATG TGTGTATGGC GGTGGAACCC CTGGCCAGAA TCATCAAGGT TATCCTGCAG
 601  TCGATGCCAG ACTTGGCTAA TGTCATGGCT CTCATCCTCT TCTTCATGCT GGTATTCTCT
 661  GTGTTTGGGG TCACGCTCTT CGGTGCATTT GTGCCCAAGC ATTTCCAGAA CATGGGGGTT
 721  GCCCTGTACA CGCTCTTCAT CTGCATCACT CAGGATGGAT GGCTGGACAT CTACACTGAC
 781  TTCCAGATGG ATGAAAGAGA GTACGCGATG GAGGTCGGGG GCGCCATCTA CTTTGCCGTC
 841  TTTATCACCC TCGGTGCCTT CATTGGTCTC AACTTGTTCG TCGTCGTGGT GACCACAAAC
 901  CTGGAACAAA TGATGAAGAC CGGCGAGGAA GAGGGACACC TGAACATAAA GTTTACTGAG
 961  ACAGAAGAGG ATGAGGACTG GACCGACGAG CTGCCACTGG TGCATTGTAC AGAGGCCCGC
1021  AAGGATACTT CCACTGTCCC CAAGGAACCA CTGGTTGGGG GCCCCCTGAG TAACCTCACA
1081  GAAAAGACCT GCGATAACTT CTGCTTGGTG CTTGAAGCAA TACAGGAGAA CTTGATGGAG
1141  TACAAAGAGA TCCGAGAGGA ACTCAACATG ATCGTGGAGG AAGTGTCCTC CATCCGGTTC
1201  AACCAGGAGC AGCAAAATGT GATCCTACAC AAGTATACCT CCAAAAGCGC CACCTTCCTA
1261  AGCGAGCCCC CAGAAGGGGC TAACAAGCAA GACTTGATCA CTGCGCTGGT CAGCAGGGAA
1321  AAGAGTAGAA CGGGTCCTGG GGATGGCTAC ACTCTCCAGT TTAGATTGTT ACTAAGCATG
1381  CTAGAGTTAC TCGGGGTGGA GCAGCCGTGG AGGCAGGAGG AGGAAGAAGG AGCAGGCTGT
1441  CCCATTAAGG TAAGCAGAGT GACAAAAAGC TCTCTTGGTT TGCAGGTGTC TGATTCTAAC
1501  ATAAACATGG TTAACAAACA CAAGTTCAGC CACTGA
```

SEQ ID: 46 INPIONCH06 amino acid sequence

```
   1  MSEKHKWWQQ VENIDITHLG PKVESLVVAL PPSWHPRRKA YELLGRHEEQ VLINRRDVME
  61  KKDAWDVQEF ITQMYIKQLL RHPAFQLLLA FLLLSNAITI ALRTNSYLGQ KHYELFSTID
 121  DIVLTILICE VLLGWLNGFW IFWKDGWNIL NFAIVFILFM GFFIKQLDMV AITYPLRVLR
 181  LVHVCMAVEP LARIIKVILQ SMPDLANVMA LILFFMLVFS VFGVTLFGAF VPKHFQNMGV
 241  ALYTLFICIT QDGWLDIYTD FQMDEREYAM EVGGAIYFAV FITLGAFIGL NLFVVVVTTN
 301  LEQMMKTGEE EGHLNIKFTE TEEDEDWTDE LPLVHCTEAR KDTSTVPKEP LVGGPLSNLT
 361  EKTCDNFCLV LEAIQENLME YKEIREELNM IVEEVSSIRF NQEQQNVILH KYTSKSATFL
 421  SEPPEGANKQ DLITALVSRE KSRTGPGDGY TLQFRLLLSM LELLGVEQPW RQEEEEGAGC
 481  PIKVSRVTKS SLGLQVSDSN INMVNKHKFS H
```

CLAIMS

1. A polypeptide, which polypeptide:

(i) comprises or consists of the amino acid sequence as recited in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO: 6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO: 20 or SEQ ID NO:22;

(ii) is a fragment thereof having the activity of a polypeptide according to (i), or having an antigenic determinant in common with a polypeptide according to (i); or

(iii) is a functional equivalent of (i) or (ii).

2. A polypeptide according to claim 1 which comprises the amino acid sequence as recited in SEQ ID NO:22.

3. A polypeptide according to claim 2 which consists of the amino acid sequence as recited in SEQ ID NO:22.

4. A polypeptide, which polypeptide:

(i) comprises or consists of the amino acid sequence as recited in SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44 or SEQ ID NO:46;

(ii) is a fragment thereof having the activity of a polypeptide of (i), or having an antigenic determinant in common with a polypeptide of (i); or

(ii) is a functional equivalent of (i) or (ii).

5.A polypeptide according to claim 4 which comprises the amino acid sequence as recited in SEQ ID NO:46.

6. A polypeptide according to claim 5 which consists of the amino acid sequence as recited in SEQ ID NO:46.

7. A polypeptide which is a functional equivalent according to claim 1 (iii) or claim 4 (iii), characterised in that it is homologous to the amino acid sequence as recited in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID

NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44 or SEQ ID NO:46 and has activity characteristic of a voltage-gated cation channel.

8. A polypeptide which is a functional equivalent according to claim 1 (iii) or claim 4 (iii), characterised in that it is homologous to the amino acid sequence as recited in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44 or SEQ ID NO:46 and has activity characteristic of a member of the CatSper family of sperm-specific cation channels.

9.A polypeptide which is a fragment or a functional equivalent according to any one of the preceding claims, which has greater than 80% sequence identity with the amino acid sequence recited in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44 or SEQ ID NO:46, or with an active fragment thereof, preferably greater than 85%, 90%, 95%, 98% or 99% sequence identity.

10. A polypeptide which is a functional equivalent as recited in any one of the preceding claims, which exhibits significant structural homology with a polypeptide having the amino acid sequence recited in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44 or SEQ ID NO:46.

11. A polypeptide which is a fragment as recited in any one of the preceding claims having an antigenic determinant in common with the polypeptide of claim 1 (i) or claim 4 (i) which consists of 7 or more (for example, 8, 10, 12, 14, 16, 18, 20 or more) amino acid residues from the amino acid sequence recited in SEQ ID NO:22 or SEQ ID NO:46.

12. A purified nucleic acid molecule which encodes a polypeptide according to any one of the preceding claims.

13. A purified nucleic acid molecule according to claim 12, which comprises the nucleic acid sequence as recited in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43 or SEQ ID NO:45 or is a redundant equivalent or fragment thereof.

14. A purified nucleic acid molecule according to claim 13 which consists of the nucleic acid sequence as recited in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43 or SEQ ID NO:45, or is a redundant equivalent or fragment thereof.

15. A purified nucleic acid molecule which hydridizes under high stringency conditions with a nucleic acid molecule according to any one of claims 12 to 14.

16. A vector comprising a nucleic acid molecule as recited in any one of claims 12 to 15.

17. A host cell transformed with a vector according to claim 16.

18. A tetrameric cation channel comprising one or more polypeptides according to any one of claims 1 to 11.

19. A tetrameric cation channel according to claim 18, additionally comprising the human CatSper1 polypeptide (genbank accession: AAL14105).

20. A tetrameric cation channel according to claim 18 or claim 19, additionally comprising a human CatSper2 polypeptide (genbank accession: AAL26490, AAL26491 or AAL26492).

21. A tetrameric cation channel according to any one of claims 18 to 20, additionally comprising the human CatSper3 polypeptide (WO200066735-A2).

22. A tetrameric cation channel according to claim 18, additionally comprising the mouse CatSper1 polypeptide (genbank accession: AAL14104).

23. A tetrameric cation channel according to claim 18 or claim 22, additionally comprising the mouse CatSper2 polypeptide (genbank accession: AAL26489).

24. A tetrameric cation channel according to any one of claims 18, 22 and 23, additionally comprising the mouse CatSper3 polypeptide (genbank accession: AK014942).

25. A tetrameric cation channel according to any one of claims 18 to 24, having the activity of sperm-specific CatSper channel.

26. A ligand which binds specifically to, and which preferably inhibits the activity of a polypeptide according to any one of claims 1 to 11 or a tetrameric cation channel according to any one of claims 18 to 25.

27. A ligand according to claim 27, which is an antibody.

28. A compound that either increases or decreases the level of expression or activity of a polypeptide according to any one of claims 1 to 11 or of a tetrameric cation channel according to any one of claims 18 to 25.

29. A compound according to claim 28 that binds to a polypeptide according to any one of claims 1 to 11 or to a tetrameric cation channel according to any one of claims 18 to 25 without inducing any of the biological effects of the polypeptide of the tetrameric cation channel.

30. A compound according to claim 28 or 29, which is a natural or modified substrate, ligand, enzyme, receptor or structural or functional mimetic.

31. A polypeptide according to any one of claims 1 to 11, a nucleic acid molecule according to any one of claims 12 to 15, a vector according to claim 16, a tetrameric cation channel according to any one of claims 18 to 25, a ligand according to claim 26 or claim 27, or a compound according to any one of claims 28 to 30, for use in therapy or diagnosis of a disease or disorder.

32. A method of diagnosing a disease in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide according to any one of claims 1 to 11, or assessing the activity of a polypeptide according to any one of claims 1 to 11, or assessing the expression or activity of a tetrameric cation channel according to any one of claims 18 to 25 in tissue from said patient and comparing said level of expression or activity to a control level, wherein a level that is different to said control level is indicative of disease.

33. A method according to claim 32 that is carried out *in vitro*.

34. A method according to claim 32 or claim 33, which comprises the steps of: (a) contacting a ligand according to claim 26 or claim 27 with a biological sample under conditions suitable for the formation of a ligand-polypeptide or ligand-tetrameric cation channel complex; and (b) detecting said complex.

35. A method according to claim 32 or claim 33, comprising the steps of:

a) contacting a sample of tissue from the patient with a nucleic acid probe under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule according to any one of claims 12 to 15 and the probe;

b) contacting a control sample with said probe under the same conditions used in step a); and

c) detecting the presence of hybrid complexes in said samples; wherein detection of levels of the hybrid complex in the patient sample that differ from levels of the hybrid complex in the control sample is indicative of a disease or disorder.

36. A method according to claim 32 or claim 33, comprising:

a) contacting a sample of nucleic acid from tissue of the patient with a nucleic acid primer under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule according to any one of claims 12 to 15 and the primer;

b) contacting a control sample with said primer under the same conditions used in step a); and

c) amplifying the sampled nucleic acid; and

d) detecting the level of amplified nucleic acid from both patient and control samples; wherein detection of levels of the amplified nucleic acid in the patient sample that differ significantly from levels of the amplified nucleic acid in the control sample is indicative of a disease or disorder.

37. A method according to claim 32 or claim 33 comprising:

a) obtaining a tissue sample from a patient being tested for a disease or disorder;

b) isolating a nucleic acid molecule according to any one of claims 12 to 15 from said

tissue sample; and

c) diagnosing the patient for disease by detecting the presence of a mutation which is associated with disease/disorder in the nucleic acid molecule as an indication of the disease.

38. The method of claim 37, further comprising amplifying the nucleic acid molecule to form an amplified product and detecting the presence or absence of a mutation in the amplified product.

39. The method of claim 37 or claim 38, wherein the presence or absence of the mutation in the patient is detected by contacting said nucleic acid molecule with a nucleic acid probe that hybridises to said nucleic acid molecule under stringent conditions to form a hybrid double-stranded molecule, the hybrid double-stranded molecule having an unhybridised portion of the nucleic acid probe strand at any portion corresponding to a mutation associated with disease; and detecting the presence or absence of an unhybridised portion of the probe strand as an indication of the presence or absence of a disease-associated mutation.

40. A method according to any one of claims 32 to 39, wherein said disease or disorder is infertility, neurological disorder, cardiovascular disorder, autoimmune disease, stroke, stroke-related disorders or other pathological condition.

41. Use of a polypeptide according to any one of claims 1 to 11 as a cation channel.

42. Use of a polypeptide according to any one of claims 1 to 11 as a calcium channel.

43. Use of a polypeptide according to any one of claims 1 to 11 or a tetrameric cation channel according to any one claims 18 to 25 as a sperm-specific CatSper channel.

44. A pharmaceutical composition comprising a polypeptide according to any one of claims 1 to 11, a nucleic acid molecule according to any one of claims 12 to 15, a vector according to claim 16, a host cell according to claim 17, a tetrameric cation channel according to any one of claims 18 to 25, a ligand according to claim 26 or claim 27, or a compound according to any one of claims 28 to 30.

45. A vaccine composition comprising a polypeptide according to any one of claims 1 to 11 or a nucleic acid molecule according to any one of claims 12 to 15.

46. A polypeptide according to any one of claims 1 to 11, a nucleic acid molecule according to any one of claims 12 to 15, a vector according to claim 16, a host cell according to claim 17, a tetrameric cation channel according to any one of claims 18 to 25, a ligand according to claim 26 or claim 27, or a compound according to any one of claims 28 to 30 or a pharmaceutical composition of claim 44, for use in the manufacture of a medicament for the treatment of a disease and disorders selected from the group consisting of infertility, neurological disorder, cardiovascular disorder, autoimmune disease, stroke, stroke-related disorders or other pathological condition and other pathological conditions.

47. A method of treating a disease in a patient, comprising administering to the patient a polypeptide according to any one of claims 1 to 11, a nucleic acid molecule according to any one of claims 12 to 15, a vector according to claim 16, a host cell according to claim 17, a tetrameric cation channel according to any one of claims 18 to 25, a ligand according to claim 26 or claim 27, or a compound according to any one of claims 28 to 30 or a pharmaceutical composition of claim 44.

48. A method according to claim 47, wherein, for diseases in which the expression of the natural gene or the activity of the polypeptide or expression or activity of the tetrameric cation channel is lower in a diseased patient or in a patient affected with a disorder when compared to the level of expression or activity in a healthy or unaffected patient, the polypeptide, nucleic acid molecule, vector, host cell, ligand, compound or composition administered to the patient is an agonist.

49. A method according to claim 47, wherein, for diseases in which the expression of the natural gene or activity of the polypeptide or expression or activity of the tetrameric cation channel is higher in a diseased patient or in a patient affected with a disorder when compared to the level of expression or activity in a healthy or unaffected patient, the polypeptide, nucleic acid molecule, vector, host cell, ligand, compound or composition administered to the patient is an antagonist.

50. A method of monitoring the therapeutic treatment of disease in a patient, comprising monitoring over a period of time the level of expression or activity of a polypeptide according to any one of claims 1 to 11, or the level of expression of a nucleic acid molecule according to any one of claims 12 to 15, or the level of expression or activity of a tetrameric cation channel according to any one of claims 18 to 25 in tissue from said patient, wherein altering said level of expression or activity over the period of time

towards a control level is indicative of regression of said disease.

51. A method for the identification of a compound that is effective in the treatment and/or diagnosis of a disease/disorder, comprising contacting a polypeptide according to any one of claims 1 to 11, or a nucleic acid molecule according to any one of claims 12 to 15, or a tetrameric cation channel according to any one of claims 18 to 25 with one or more compounds suspected of possessing binding affinity for said polypeptide, nucleic acid molecule or tetrameric cation channel, and selecting a compound that binds specifically to said polypeptide, nucleic acid molecule or tetrameric cation channel.

52. A kit useful for diagnosing disease comprising a first container containing a nucleic acid probe that hybridises under stringent conditions with a nucleic acid molecule according to any one of claims 12 to 15; a second container containing primers useful for amplifying said nucleic acid molecule; and instructions for using the probe and primers for facilitating the diagnosis of disease.

53. The kit of claim 52, further comprising a third container holding an agent for digesting unhybridised RNA.

54. A kit comprising an array of nucleic acid molecules, at least one of which is a nucleic acid molecule according to any one of claims 12 to 15.

55. A kit comprising one or more antibodies that bind to a polypeptide as recited in any one of claims 1 to 11 or to a tetrameric cation channel as recited in any one of claims 18 to 25, and a reagent useful for the detection of a binding reaction between said antibody and said polypeptide/tetrameric cation channel.

56. A transgenic or knockout non-human animal that has been transformed to express higher, lower or absent levels of a polypeptide according to any one of claims 1 to 11.

57. A method for screening for a compound effective to treat a disease/disorder, by contacting a non-human transgenic animal according to claim 56 with a candidate compound and determining the effect of the compound on the disease/disorder of the animal.

58. A polypeptide according to any one of claims 1 to 11, a nucleic acid molecule according to any one of claims 12 to 15, a vector according to claim 16, a host cell according to claim 17, a tetrameric cation channel according to any one of claims 18 to 25, a ligand according to claim 26 or claim 27, or a compound according to any one of claims 28 to 30, or a pharmaceutical composition according to claim 44, for use as a contraceptive

agent.

59. Claim 58, wherein the contraceptive agent is a non-hormonal contraceptive agent.

Figure 1: Top 15 blastp search results against NCBI non-redundant database using the full length INPIONCH05 polypeptide sequence

Sequences producing significant alignments:	Score (bits)	E Value
gi|15614064|ref|NP_242367.1| (NC_002570) BH1501~unknown con...	102	1e-20
gi|4755118|gb|AAD23600.2| (AF134216) putative sodium channe...	92	9e-18
gi|12658323|gb|AAK01090.1| (AF312365) putative BSC1 sodium ...	92	1e-17
gi|7293235|gb|AAF48617.1| (AE003502) para gene product [Dro...	91	3e-17
gi|18489453|ref|XP_082886.1| (XM_082886) paralytic [Drosoph...	91	3e-17
gi|2119597|pir||I54323 sodium channel alpha subunit - human...	91	4e-17
gi|1842214|gb|AAB47605.1| (U38814) voltage-sensitive sodium...	91	4e-17
gi|1783373|gb|AAB47604.1| (U38813) voltage-sensitive sodium...	91	4e-17
gi|7441675|pir||S72458 sodium channel protein para-type alp...	91	4e-17
gi|7522168|pir||T31340 voltage-gated sodium channel homolog...	90	7e-17
gi|16076814|gb|AAL14104.1|AF407332_1 (AF407332) sperm ion c...	89	9e-17
gi|9507065|ref|NP_062138.1| (NM_019265) sodium channel, vol...	89	9e-17

Figure 2: Alignment generated by BLAST between the full-length INPIONCH05 polypeptide sequence and the sequence of the closest known protein, (L01983) sodium channel alpha subunit from *Homo sapiens*.

```
>gi|2119597|pir||I54323 sodium channel alpha subunit - human
 gi|908809|gb|AAA75557.1| (L01983) sodium channel alpha subunit [Homo sapiens]
          Length = 1835

 Score = 90.5 bits (223), Expect = 4e-17
 Identities = 66/281 (23%), Positives = 128/281 (45%), Gaps = 51/281 (18%)

Query: 87   QLLRHPAFQLLLALLLVINAITIALRTNSYLDQKHY--ELFSTIDDIVLTILLCEVLLGW 144
            +++ H  F+  +  ++++++  +A   + Y++Q+     +    D +   I + E+LL W
Sbjct: 1026 KIVEHNWFETFIVFMILLSSGALAFE-DIYIEQRKVIRTILEYADKVFTYIFIMEMLLKW 1084

Query: 145  LN-GFWIFWKDGWNILNFIIVFILLLRFFINEINIPSIN-----YTLRALRLVHVCMAVE 198
            +  GF +++ + W  L+F+IV + ++     N +     +        TLRALR +       E
Sbjct: 1085 VAYGFKVYFTNAWCWLDFLIVDVSIISLVANWLGYSELGPIKSLRTLRALRPLRALSRFE 1144

Query: 199  PLARIIRVILQSVPDMANIMVLILFFMLVFSVFGVTLF-GAF------------------ 239
             +  ++  +L ++P + N+++ L F L+FS+ GV LF G F
Sbjct: 1145 GMRVVVNALLGAIPSIMNVLLVCLIFWLIFSIMGVNLFAGKFYYCINTTTSERFDISEVN 1204

Query: 240  ------------------VPKHFQNIQVALYTLFICITQDGWVDIY-----SDFQTEKRE 276
                              V  ++ N+ +   +L    T  GW+DI       S  + E+ +
Sbjct: 1205 NKSECESLMHTGQVRWLNVKVNYDNVGLGYLSLLQVATFKGWMDIMYAAVDSREKEEQPQ 1264

Query: 277  YAMEIGGAIYFTIFITIGAFIGINLFVIVVTTNLEQMMKAG 317
            Y + +   +YF IFI  G+F  +NLF+ V+  N  Q  K G
Sbjct: 1265 YEVNLYMYLYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKG 1305
```

Figure 3: Results from BLAST against NCBI non-redundant database using the masked full length INPIONCH05 sequence. By 'masked' it is meant that where there exists a long row of hydrophobic residues, these have been replaced ('masked') by a row of Xs.

Sequences producing significant alignments:	Score (bits)	E Value
gi|15614064|ref|NP_242367.1| (NC_002570) BH1501~unknown con...	66	7e-10
gi|18578174|ref|XP_083858.1| (XM_083858) cation channel of ...	56	1e-06
gi|16596698|ref|NP_444282.1| (NM_053054) cation channel of ...	55	1e-06
gi|7522168|pir||T31340 voltage-gated sodium channel homolog...	54	3e-06
gi|16076814|gb|AAL14104.1|AF407332_1 (AF407332) sperm ion c...	54	4e-06
gi|9507065|ref|NP_062138.1| (NM_019265) sodium channel, vol...	53	1e-05
gi|2119597|pir||I54323 sodium channel alpha subunit - human...	52	1e-05
gi|4755118|gb|AAD23600.2| (AF134216) putative sodium channe...	51	3e-05
gi|1352087|sp|P35499|CIN4_HUMAN Sodium channel protein, ske...	50	5e-05
gi|4506807|ref|NP_000325.1| (NM_000334) sodium channel, vol...	50	5e-05

Figure 4: Alignment generated by BLAST between the masked INPIONCH05 polypeptide sequence and the sequence of the closest known protein, (XM_083858.1) cation channel of sperm from *Homo sapiens*

```
>gi|18578174|ref|XP_083858.1|  [LocusLink info] (XM_083858) cation channel of sperm
[Homo sapiens]
          Length = 763

 Score = 55.8 bits (133), Expect = 1e-06
 Identities = 33/93 (35%), Positives = 50/93 (53%), Gaps = 3/93 (3%)

Query: 184 TLRALRLVHVCMAVEPLARIIRV---ILQSVPDMANIMVLILFFMLVFSVFGVTLFGAFV 240
           +LRALR + V    +  L  +  V    + QS+P +A I++L+    + +FS      LF
Sbjct: 550 SLRALRAIRVLRRLSFLTSVQEVTGTLGQSLPSIAAILILMFTCLFLFSAVLRALFRKSD 609

Query: 241 PKHFQNIQVALYTLFICITQDGWVDIYSDFQTE 273
           PK FQNI    ++TLF  +T D W  IY D + +
Sbjct: 610 PKRFQNIFTTIFTLFTLLTLDDWSLIYMDSRAQ 642
```

Figure 5: Conserved domain search results: Position Specific Scoring Matrices (PSSMs) producing significant alignments to the amino acid sequence of INPIONCH05

```
PSSMs producing significant alignments:                    Score(bits) value

gnl|Pfam|pfam00520 ion_trans, Ion transport protein.        96.7        3e-21

gnl|Pfam|pfam00520 , ion_trans, Ion transport protein. This family contains Sodium,
Potassium, Calcium ion channels. This family is 6 transmembrane helices in which
the last two helices flank a loop which determines ion selectivity. In some sub-
families (e.g. Na channels) the domain is repeated four times, whereas in others
(e.g. K channels) the protein forms as a tetramer in the membrane. A bacterial
structure of the protein is known for the last two helices but is not the Pfam
family due to it lacking the first four helices

 CD-Length = 191 residues, 100.0% aligned
            Score = 96.7 bits (239), Expect = 3e-21

Query:  124  LFSTIDDIVLTILLCEVLLGW--LNGFWIFWKDGWNILNFIIVFILLLRFFINEINIPSI 181
Sbjct:  1    ILEILDYVFTVIFTLEMLLKFIALGFKLKYLRSPWNILDFLIVLPSLIDLILFLSGGGSV 60

Query:  182  NYTLRALRLVHVCMAVEPLARIIRVILQSVPDMANIMVLILFFMLVFSVFGVTLFGA--- 238
Sbjct:  61   LRLLRLLRLLRLLRRLEGLRTLLQSLGRSLKSLLNLLLLLLLLLFIFAIIGVQLFGGEFN 120

Query:  239  ---------FVPKHFQNIQVALYTLFICITQDGWVDIYSDFQTEKREYAMEIGGAIYFTI 289
Sbjct:  121  KCCDGVNPINGNSNFDSFGEAFYWLFRTLTTEGWGDIMPD------TLDAPVLGKIFFVI 174

Query:  290  FITIGAFIGINLFVIVV 306
Sbjct:  175  FIILGGLLLLNLFIAVI 191
```

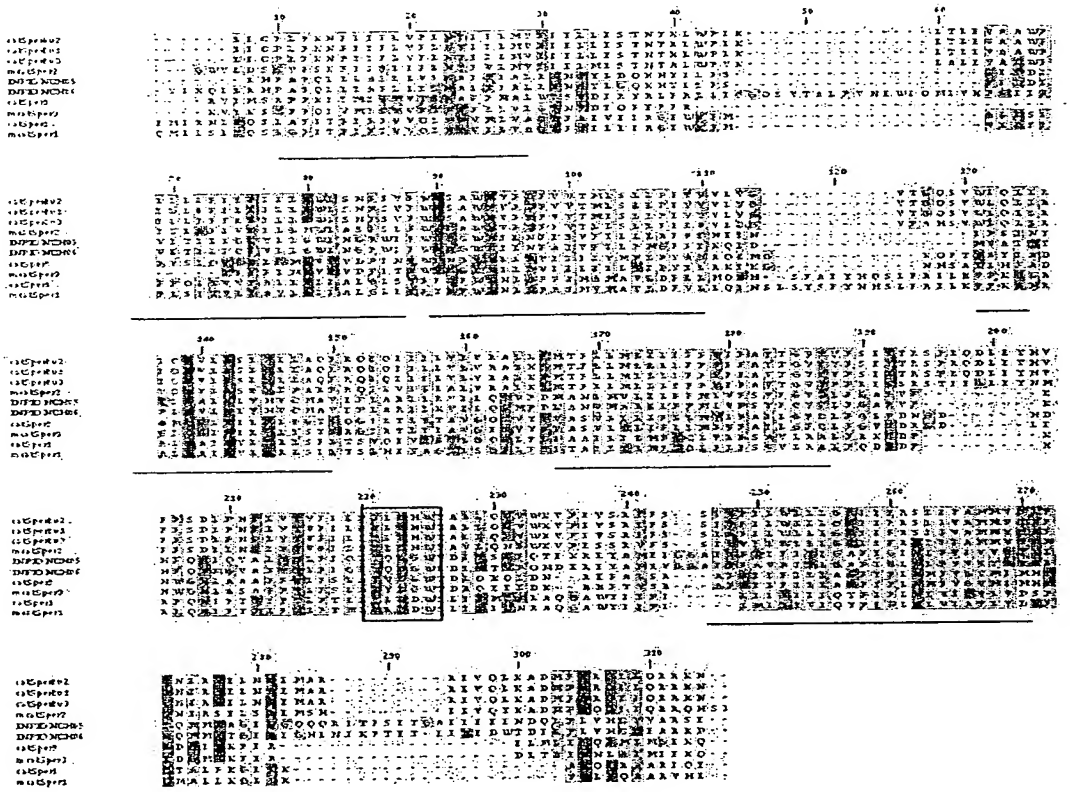
Figure 6: Interpro scan results for INPIONCH05

```
InterPro
IPR000636 Cation channel, non-ligand gated  (Family)
PROFILE: PS50266  CATION_CHANNEL_TM   [92-310]T
PFAM: PF00520  ion_trans   [124-306]T

InterPro
IPR002111 Cation (not K+) channel, TM region  (Domain)
PROFILE: PS50272  CATION_CHANNEL_TRPL   [132-310]T

InterPro
IPR001682 Ca2+/Na+ channel, pore region  (Family)
PROFILE: PS50273  CHANNEL_PORE_CA_NA   [180-310]T
```

Figure 7: Ion Trans Domain Multiple Alignment between CatSper family members showing transmembrane regions, and TxDxW ion_selectivity region boxed



Key:

catSper1	Genbank accession:AAL14105	human catSper1
mcatSper1	Genbank accession:AAL14104	mouse catSper1
catSper2v1	Genbank accession:AAL26490	human catSper2 variant 1
catSper2v2	Genbank accession:AAL26491	human catSper2 variant 2
catSper2v3	Genbank accession:AAL26492	human catSper2 variant 3
mcatSper2	Genbank accession:AAL26489	mouse catSper2
catSper3	Patent:WO200066735-A2	human catSper3
mcatSper3	Genbank accession: AK014942	mouse catSper3
INPIONCH05		human catSper4
INPIONCH06		mouse catSper4

Figure 8: INPIONCH05 topology prediction using the TMHMM programme for prediction of helices in proteins.

TMHMM2.0	inside	1	92
TMHMM2.0	TMhelix	93	112
TMHMM2.0	outside	113	126
TMHMM2.0	TMhelix	127	149
TMHMM2.0	inside	150	155
TMHMM2.0	TMhelix	156	173
TMHMM2.0	outside	174	217
TMHMM2.0	TMhelix	218	240
TMHMM2.0	inside	241	246
TMHMM2.0	TMhelix	247	264
TMHMM2.0	outside	265	283
TMHMM2.0	TMhelix	284	306
TMHMM2.0	inside	307	522

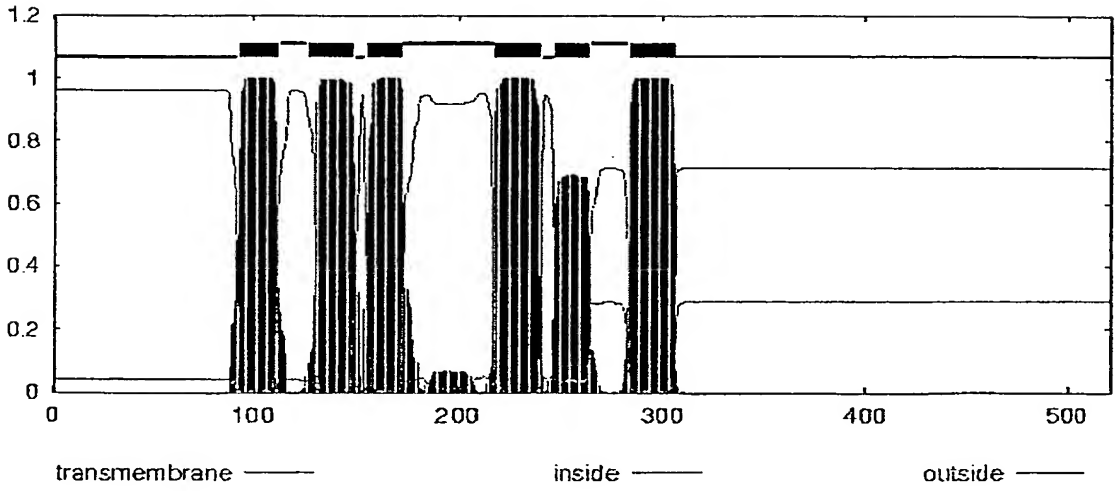
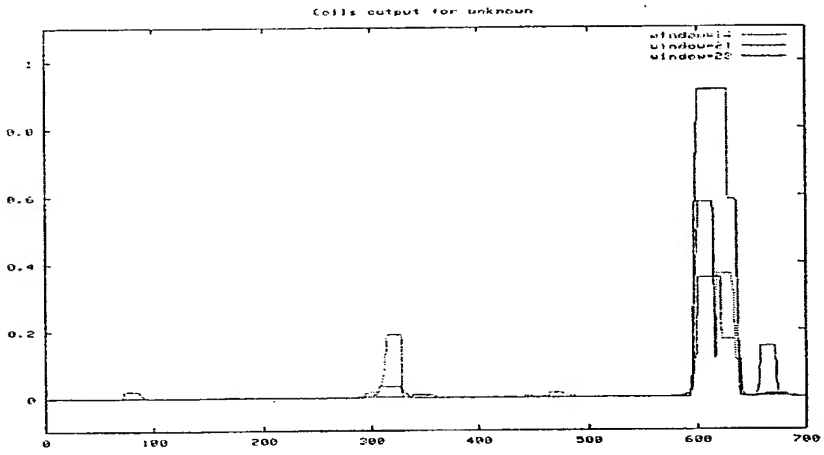


Figure 9: Coiled-coil prediction from Lupas Coiled coils prediction program

CatSper1 coiled coil prediction



CatSper2 coiled coil prediction

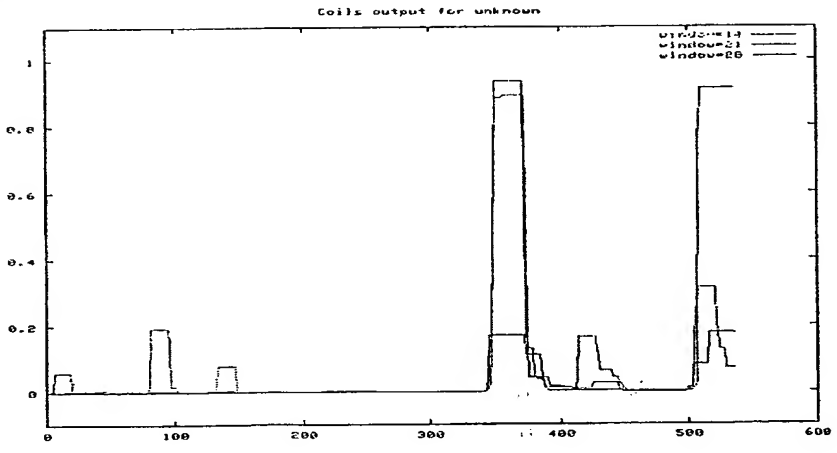
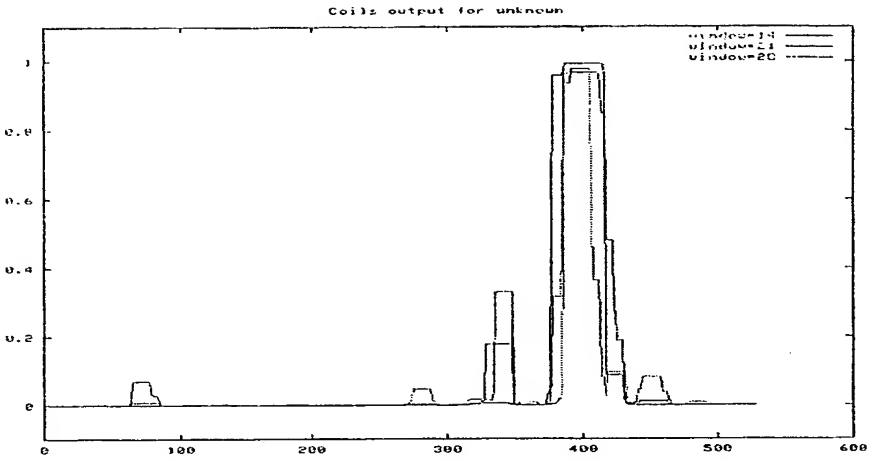


Figure 9 continued...

INPIONCH05 coiled coil prediction



CatSper3 coiled coil prediction

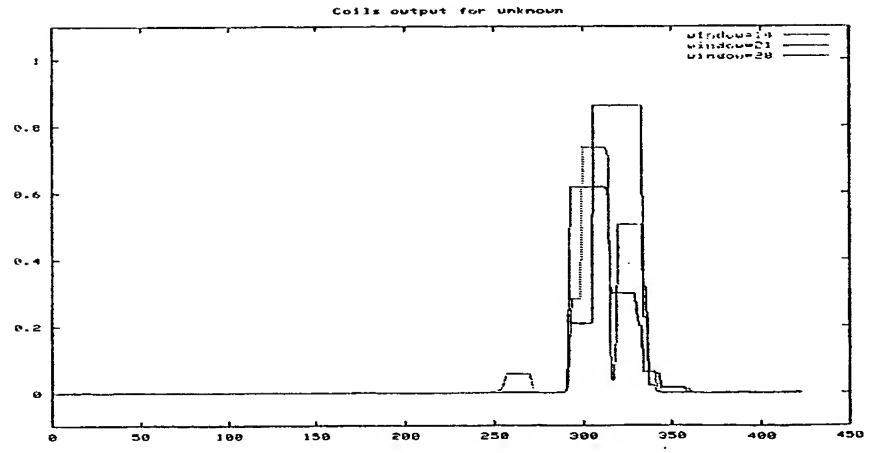


Figure 12: Sequences of CatSper3

CatSper3 Lexicon genetics human sequence WO200066735-A2

```
  1  MSQHRHQRHS RVISSSPVDT TSVGFCPTFK KFKRNDDECR AFVKRVIMSR FFKIIMISTV
 61  TSNAFFMALW TSYDIRYRLF RLLEGGCSVT ALPVNEWICM IVKFSEIFFV SICTSELSMK
121  VYVDPINYWK NGYNLLDVII IIVMFLPYAL RQLMGKQFTY LYIADGMQSL RILKLIGYSQ
181  GIRTLITAVG QTVYTVASVL LLLFLLMYIF AILGFCLFGS PDNGDHDNWG NLAAAFFTLF
241  SLATVDGWTD LQKQLDNREF ALSRAFTIIF ILLASFIFLN MFVGVMIMHT EDSIRKFERE
301  LMLEQQEMLM GEKQVILQRQ QEEISRLMHI QKNADCTSFS ELVENFKKTL SHTDPMVLDD
361  FGTSLPFIDI YFSTLDYQDT TVHKLQELYY EIVHVLSLML EDLPQEKPQS LEKVDEK
```

CatSper3 Lexicon genetics mouse sequence WO200066735-A2

```
  1  MSQHFHHNFV RVKSGSLFAT ASEALQARLS KIKRKDKECQ AYFRKVIKST FFQIVMITTV
 61  TTNSFLLVLG TNYDIQFEFF RTFEFLMKVY VDPITYWKDG YNILDVIILI ILTIPYLLRK
121  IKGNHSAYLH FADGIQSLRI LKLISYSRGI RTLIIAVGET VYTVASVLTL LFLLMFVFAI
181  LGFCLFGVTD RGDLENWGNL ASAFFTLFSL ATVDGWTDLQ EELDKRKFTV SRAFTILFIL
241  LASFIFLNMF VGVMIMHTED SMKKFERDLT LERNLAIMEE KQIILKRQQE EVNRLMNTQK
301  SGSMNFIDMV EGFKKTLRHT DPMVLDDFST SLSFIDIYLV TLDNQDVIVS KLQELYCEIV
361  NVLSLMLEDM PKESSSSLSG LS
```

Figure 13: Sequences of CatSper2 human splice variants

CatSper2v2 human splice variant AAL26491

```
  1  MAAYQQEEQM QLPRADAIRS RLIDTFSLIE HLQGLSQAVP RHTIRELLDP SRQKKLVLGD
 61  QHQLVRFSIK PQRIEQISHA QRLLSRLHVR CSQRPPLSLW AGWVLECPLF KNFIIFLVFL
121  NTIILMVEIE LLESTNTKLW PLKLTLEVAA WFILLIFILE ILLKWLSNFS VFWKSAWNVF
181  DFVVTMLSLL PEVVVLVGVT GQSVWLQLLR ICRVLRSLKL LAQFRQIQII ILVLVRALKS
241  MTFLLMLLLI FFYIFAVTGV YVFSEYTRSP RQDLEYHVFF SDLPNSLVTV FILFTLDHWY
301  ALLQDVWKVP EVSRIFSSIY FILWLLLGSI IFRSIIVAMM VTNFQNIRKE LNEEMARREV
361  QLKADMFKRQ IIQRRKNMSH EALTSSHSKI EDSSRGASQQ RESLDLSEVS EVESNYGATE
421  EDLITSASKT EETLSKKREY QSSSCVSSTS SSYSSSSESR FSESIGRLDW ETLVHENLPG
481  LMEMDQDDRV WFRDSLFRYF ELLEKLQYNL EERKKLQEFA VQALMNLEDK
```

CatSper2v3 human splice variant AAL26492

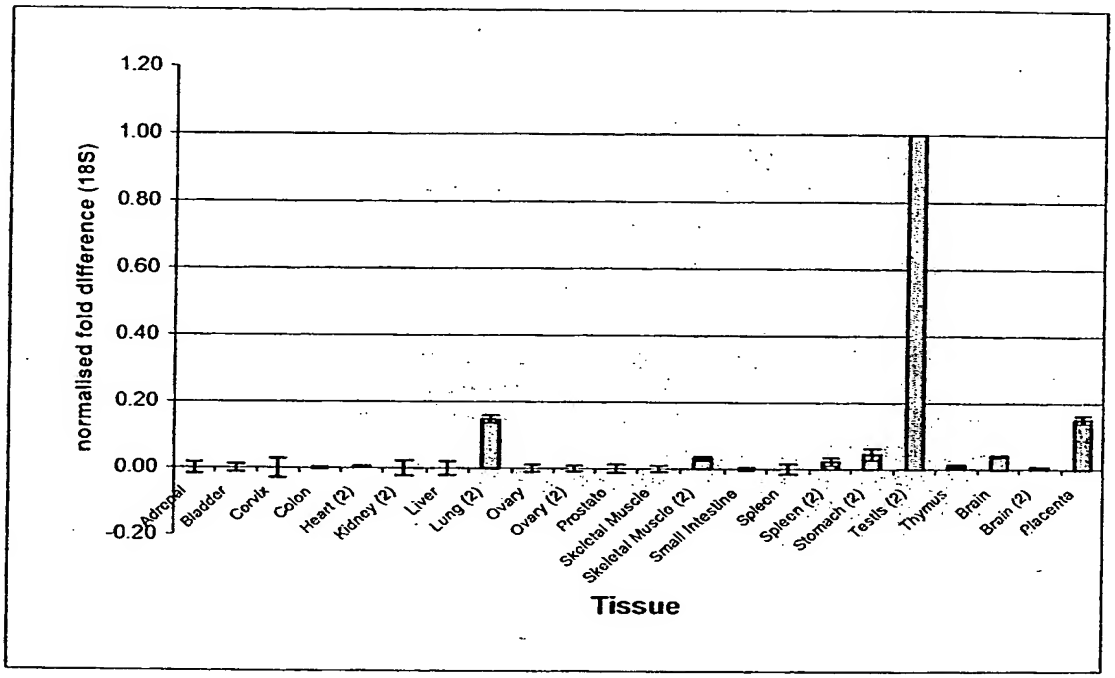
```
  1  MAAYQQEEQM QLPRADAIRS RLIDTFSLIE HLQGLSQAVP RHTIRELLDP SRQKKLVLGD
 61  QHQLVRFSIK PQRIEQISHA QRLLSRLHVR CSQRPPLSLW AGWVLECPLF KNFIIFLVFL
121  NTIILMVEIE LLESTNTKLW PLKLTLEVAA WFILLIFILE ILLKWLSNFS VFWKSAWNVF
181  DFVVTMLSLL PEVVVLVGVT GQSVWLQLLR ICRVLRSLKL LAQFRQIQII ILVLVRALKS
241  MTFLLMLLLI FFYIFAVTGV YVFSEYTRSP RQDLEYHVFF SDLPNSLVTV FILFTLDHWY
301  ALLQDVWKVP EVSRIFSSIY FILWLLLGSI IFRSIIVAMM VTNFQNIRKE LNEEMARREV
361  QLKADMFKRQ IIQRRKNMSH EALTSSHSKI EDRSFGLGDS CARKSARANG NGSG
```

Figure 14: ENSEMBL predicted amino acid transcript: ENST00000294438

Human assembly version 4.28.1 12 March 2002 derived from ENSG00000162449 predicted gene

```
  1  STIHESYGRP EEQVLINRQE ITNKADAWDM QEFITHMYIK QLLRHPAFQL LLALLLVINA
 61  ITIALRTNSY LDQKHYELFS TIDDIVLTIL LCEVLLGWLN GFWIFWKDGW NILNFIIVFI
121  LLLRFFINEI NIPSINYTLR ALRLVHVCMA VEPLARIIRV ILQSVPDMAN IMVLILFFML
```

Figure 15: Normalised expression of INPIONCH05 in human tissue.



INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 03/02270

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/705 C12N5/10 C12N15/12 C12Q1/68 A01K67/027
A61K38/02 A61K39/00 G01N33/48 G01N33/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K A01K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EMBL, MEDLINE, WPI Data, EPO-Internal, SEQUENCE SEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 'Online! standard; RNA; EST; 605 BP, 20 October 2001 (2001-10-20) ARAKAWA T. ET AL.: "Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone: 4933412L05, 5' end partial sequence" Database accession no. BB617038 XP002253118 the whole document --- -/--	12-17

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 September 2003	24/09/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Kools, P

Form PCT/ISA/210 (second sheet) (July 1992)

THIS PAGE BLANK (USP